

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using BW model

Run on: October 4, 2005, 18:34:55 ; Search time 166 Seconds
(without alignments)
146.783 Million cell updates/sec

Title: US-10-669-175-1
Perfect score: 327
Sequence: 1 ALENLVVNNASAGTHGIL.....LLGLWPLLLLLLALPQRAYA 63

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	283	86.5	283	AAR29852	Aar29852 HCV NS2-N
2	283	86.5	480	AAR34002	Aar34002 BK E2/NS1
3	283	86.5	3010	AAR20111	Aar20111 Non-A, no
4	283	86.5	3010	AAR20091	Aar20091 Non-A, no
5	283	86.5	3010	AAY06423	Aay06423 Non-A, no
6	281	85.9	1026	ADJ82984	Adj82984 Adenovect
7	281	85.9	2307	AAY70064	Aay70064 Recombina
8	281	85.9	2307	AAY70065	Aay70065 Recombina
9	281	85.9	2307	AAY70066	Aay70066 Recombina
10	281	85.9	3010	ABG32458	Abg32458 Hepatitis
11	281	85.9	3010	ABG32459	Abg32459 Hepatitis
12	281	85.9	3010	ABG32451	Abg32451 Hepatitis
13	281	85.9	3010	ABG32455	Abg32455 Hepatitis
14	281	85.9	3010	ABG32457	Abg32457 Hepatitis
15	281	85.9	3010	ABG32460	Abg32460 Hepatitis
16	281	85.9	3010	ABG32453	Abg32453 Hepatitis
17	281	85.9	3010	ABG32461	Abg32461 Hepatitis
18	281	85.9	3010	ABG32454	Abg32454 Hepatitis
19	281	85.9	3010	ABG32452	Abg32452 Hepatitis
20	281	85.9	3011	ABG32456	Abg32456 Hepatitis
21	279	85.3	3010	AAR30616	Aar30616 Polyperi
22	278	85.0	3010	AAE20477	Aae20477 HCV-S1 fu
23	277	84.7	3011	AAR34468	Aar34468 Encoded b
24	275	84.1	258	AAW41735	Aaw41735 Hepatitis
25	275	84.1	259	AAR25857	Aar25857 HCV poly

26	275	84.1	480	2	AAR34000
27	275	84.1	980	2	AAR53921
28	275	84.1	1010	2	AAR88311
29	275	84.1	3010	2	AAR68622
30	275	84.1	3010	2	AAR68864
31	275	84.1	3010	2	AAR82694
32	275	84.1	3010	7	ADF88597
33	274	83.8	283	2	AAR29853
34	274	83.8	283	2	AAR29851
35	274	83.8	845	2	AAW68466
36	274	83.8	1031	2	AAR54067
37	274	83.8	1031	2	AAR98362
38	274	83.8	1188	2	AAR29870
39	274	83.8	3010	2	AAR34580
40	273	83.5	1051	2	AAR54066
41	273	83.5	1051	2	AAR98361
42	273	83.5	3010	2	AAW98022
43	273	83.5	3010	4	AAW59174
44	273	83.5	3010	4	AAW31170
45	273	83.5	3010	8	ADO36227

ALIGNMENTS

RESULT 1		AAR29852 standard; protein; 283 AA.	
ID	AAR29852		
AC	AAR29852;		
XX			
DT	25-MAR-2003 (revised)		
DT	26-APR-1993 (first entry)		
XX			
DE	HCV NS2-NS4 peptide WX25-2.		
XX			
KW	Clone; polypeptide; NS2-NS4; Hepatitis C virus; HCV; serum; HC;		
KW	transcriptase; cDNA; primer; allele.		
OS	Hepatitis C virus.		
XX			
PN	EP518313-A2.		
XX			
PD	16-DEC-1992.		
XX			
PF	11-JUN-1992; 92EP-00109812.		
XX			
PR	11-JUN-1991; 91JP-00139268.		
PR	12-JUL-1991; 91JP-00172794.		
PR	07-OCT-1991; 91JP-00287008.		
PR	16-DEC-1991; 91JP-00332329.		
PR	20-APR-1992; 92JP-00099957.		
PA	(MITU) MITSUBISHI KASEI CORP.		
XX			
PI	Seki M, Honda Y, Takahashi K, Murakami T, Teranishi Y, Hayashi N;		
XX	WPI; 1992-417213/51.		
DR	N-PSDB; AAQ32483.		
XX			
PT	New hepatitis C virus gene and its encoded protein - used for diagnosing		
PT	and vaccinating against hepatitis C virus infections.		
XX			
PS	Disclosure: Page 147-49; 305PP; English.		
XX			
CC	The sequences given in AAR29852-70 are encoded by various clones which		
CC	were used in the isolation of the NS2-NS4 regions of the Hepatitis C		
CC	virus (HCV) gene of the invention (see also AAR29660, AAR29559-60 and		
CC	AAR29843-51). These RNA sequences were isolated from the serum of a		
CC	patient suffering from hepatitis C (HC). The isolated RNA sequences were		
CC	converted into cDNA using transcriptase in the presence of one of the		
CC	primer sequences given in AAQ32578-79. The sequences were then amplified		
CC	using primer pairs. The cDNA sequences isolated represent different		

Aar34000	HCV-J E2/
Aar53321	HCV fusio
Aar88311	HCV pepti
Aar68622	HCV prote
Aar68864	Hepatitis
Aar82694	Partial H
Adf88597	Hepatitis
Aar29853	HCV NS2-N
Aar29851	HCV NS2-N
AAW68466	Protein e
Aar54067	Non-A, no
Aar98362	5'UTR/COR
Aar29870	HCV NS2-N
Aar34580	Human hep
Aar54066	Non-A, no
Aar98361	5'UTR/COR
AAW98022	Infectiou
AAW59174	Protein e
AAW31170	Amino aci
ADO36227	Hepatitis

CC alleles of the same region of the HCV gene. Sequence comparisons of these
 CC clones showed that it is possible for a patient to carry more than one
 CC HCV strain at one time. See also AAQ32436. (Updated on 25-MAR-2003 to
 CC correct PN field.)
 XX

SO Sequence 283 AA;

Query Match 86.5%; Score 283; DB 2; Length 283;
 Best Local Similarity 85.7%; Pred. No. 9.6e-29;
 Matches 54; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 ALENLVNNAASAGHGLIMPLVFCAMVYKGLVPGATYSLGLMPLLLLLALPQR 60
 DB 16 ALENLVNNAASAGHGLISFLVFCAMVYIKGRLVPGATYALYGVWPLLLLLALPQR 75

QY 61 AYA 63
 DB 76 AYA 78

RESULT 2
 AAR34002

ID AAR34002 standard; protein; 480 AA.

AC AAR34002;

DT 25-MAR-2003 (revised)

DT 26-JUL-1993 (first entry)

XX BK E2/NS1 protein.

XX Polymerase chain reaction; PCR; amplify; primer; hepatitis C virus; HCV;
 KW asymptomatic; chronically infected; epitope; viral isolate; domain;
 KW immunological; cross-reactive; envelope protein; vaccine;

KW gp53 (BVDV)/gp55; hog cholera virus; pestivirus; NS1; flavivirus.

OS Synthetic.

PN WO9306126-A1.

PD 01-APR-1993.

PF 11-SEP-1992; 92MO-US007683.

PR 13-SEP-1991; 91US-00759575.

PA (CHIR) CHIRON CORP.

PI Weiner AJ, Houghton M;

DR WPI; 1993-117468/14.

PT Immuno-reactive hepatitis C virus polypeptide compans. - contg. at least
 PT 2 sequences from the first variable domain of distinct HCV isolates.

XX Disclosure; Fig 3; 106pp; English.

CC The sequences given in AAR33992-002 represent a portion of the E2/NS1
 CC protein encoded by group I and group II HCV isolates, from amino acid 370
 CC -850. E2/NS1 are viral envelope proteins and are of immunogenic interest.
 CC E2/NS1 contains an N-terminal hypervariable domain of about 30 amino
 CC acids which shows large variation between nearly all isolates. This is an
 CC important immunoreactive domain. This putative envelope glycoprotein
 CC E2/NS1 may correspond to the gp53(BVDV)/gp55 (hog cholera virus) envelope
 CC polypeptide of the pestiviruses and the NS1 of the flaviviruses, both of
 CC which confer protective immunity in hosts vaccinated with these
 CC polypeptides. It has been discovered that a number of important HCV
 CC epitopes vary among viral isolates and that these epitopes can be mapped
 CC to specific domains. This meant that immunologically cross-reactive
 CC polypeptides which focus on variable rather than constant domains can be
 CC produced. See also AAQ39134-48 and AAR33982-91. (Updated on 25-MAR-2003
 CC to correct PN field.)
 XX

SO Sequence 480 AA;

Query Match 86.5%; Score 283; DB 2; Length 480;
 Best Local Similarity 85.7%; Pred. No. 1.8e-28;
 Matches 54; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 ALENLVNNAASAGHGLIMPLVFCAMVYKGLVPGATYSLGLMPLLLLLALPQR 60
 DB 378 ALENLVNNSASVAGAHGLISFLVFCAMVYIKGRLVPGATYALYGVWPLLLLLALPQR 437

QY 61 AYA 63
 DB 438 AYA 440

RESULT 3
 AAR20111

ID AAR20111 standard; protein; 3010 AA.

AC AAR20111;

DT 25-MAR-2003 (revised)

DT 01-MAY-1992 (first entry)

XX Non-A, non-B viral genome product.

XX NANBV; vaccine; immunodiagnosis; antigen; antibody.

OS Non-A.

XX non-B hepatitis virus.

XX Key Location/Qualifiers

FT Protein

FT 1. .115

FT /label= C

FT /note= "core protein"

FT 116. .191

FT /label= M

FT /note= "matrix protein"

FT 192. .389

FT /label= E

FT /note= "envelope protein"

FT 390. .729

FT /label= NS1

FT 730. .1006

FT /label= NS2

FT 1007. .1614

FT /label= NS3

FT 1615. .1862

FT /label= NS4a

FT 1863. .2012

FT /label= NS4b

FT 2013. .3010

FT /label= NS5

PN EP463848-A.

PD 02-JAN-1992.

PF 25-JUN-1991; 91EP-00305717.

PR 25-JUN-1990; 90JP-00167466.

PR 31-AUG-1990; 90JP-00230921.

PR 09-NOV-1990; 90JP-00305605.

PR 28-DEC-1990; 90US-00635451.

PR 08-MAY-1991; 91JP-00132090.

PR 14-MAY-1991; 91JP-00138493.

XX (OSAU) UNITV OSAKA.

PA (REMI-) RES FOUND MICROBIAL.

DR WPI; 1992-009412/02.
 DR N-PSDB; AAQ20268.
 XX

PT Non-A, non-B hepatitis virus (NANBV) particles - as vaccinees, immuno-
PT diagnostics and screening agents for NANBV, and to remove NANBV from
PT blood.
XX
XX Disclosure; Fig 2; 89pp; English.
XX
XX The sequence (SEQ ID NO 1) was deduced from several overlapping from a
CC library prep. from NANBV RNA. Antigenic polypeptides from the sequence
CC can be used as immunoassay reagents, for screening donated blood, and as
CC immunogens for vaccine prodn. Antibodies raised to the peptides can be
CC used in immunoassays to detect or quantify NANBV antigens in liver tissue
CC and blood. Preferred poly- peptides include residues 1-30, -115, or 2012;
CC 47-77; 116-191; 192-207 or -263; 287-300; 293-330; 390-
CC 729; 730-1005; 1006-1614; 1384-1414; 1615-1862; 1737-1767; 1863-2012;
CC and 2013-3010. The sequence is also disclosed in EP-464287 (SEQ ID NO 1).
CC See AAR20091 for details of this specification. (Updated on 25-MAR-2003
CC to correct PD field.) (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 3010 AA;
SQ

Query Match 86.5%; Score 283; DB 2; Length 3010;
Best Local Similarity 85.7%; Pred. No. 1.4e-27;
Matches 54; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 1 ALENLVNMAASAGTGILMFLVFCAMVYKRLVGATYSLLGLMPLLLLLALPQR 60
DB 747 ALENLVNMAASAGTGILMFLVFCAMVYKRLVGATYSLLGLMPLLLLLALPQR 806
OY 61 AYA 63
DB 807 AYA 809

RESULT 4
AAR20091
ID AAR20091 standard; protein; 3010 AA.
XX
XX AAR20091;
AC
XX
DT 25-MAR-2003 (revised)
DT 01-MAY-1992 (first entry)
XX
XX Non-A, non-B viral genome product.
DE
XX NANBV; vaccine; immunodiagnosis; antigen; antibody.
KW
XX
XX Non-A.
OS non-B hepatitis virus.
OS
XX
XX
FH Key Location/Qualifiers
FT 1. .115
FT /label= C
FT /note= "core protein"
FT 116. .191
FT /label= M
FT /note= "matrix protein"
FT 192. .389
FT /label= E
FT /note= "envelope protein"
FT 390. .729
FT /label= NS1
FT 730. .1006
FT /label= NS2
FT 1007. .1614
FT /label= NS3
FT 1615. .1862
FT /label= NS4a
FT 1863. .2012
FT /label= NS4b
FT 2013. .3010
FT /label= NS5
XX
XX EP464287-A.

XX
PD 08-JAN-1992.
XX
XX 28-DEC-1990; 90EP-00314371.
PF
XX 25-JUN-1990; 90JP-00167466.
XX
PR 31-AUG-1990; 90JP-00230921.
PR 09-NOV-1990; 90JP-00305605.
PR 17-JUN-1991; 91EP-00401604.
XX
XX (OSAU) UNIV OSAKA.
PA
XX
XX WPI; 1992-009617/02.
DR N-PSDB; AAQ21829.
XX
XX New DNA from non-A, non-B hepatitis virus - and derived antigenic
PT polypeptide(s) useful for diagnostics, blood screening and in vaccines.
XX
XX Claim 3; Fig 2; 89pp; English.
XX
XX The sequence was deduced from several overlapping "BK" cDNA clones obt.
CC by "gene walking" using a cDNA clone isolated from a library prep. from
CC NANBV RNA. Antigenic polypeptides from the sequence can be used as
CC immunoassay reagents, for screening donated blood, and as immunogens for
CC vaccine prodn. Antibodies raised to the peptides can be used in
CC immunoassays to detect or quantify NANBV antigens in liver tissue and
CC blood. Preferred polypeptides are include residues 1-30, -115, or 2012;
CC 47-77; 116-191; 192-207 or -298; 230-238 or -263; 287-300; 293-330; 390-
CC 729; 730-1005; 1006-1614; 1384-1414; 1615-1862; 1737-1767; 1863-2012;
CC and 2013-3010. The sequence is also disclosed in EP-463848 (SEQ ID NO 1)
CC in which a virus particle contg. antigens encoded by the sequence is
CC claimed. See AAR20111 for details of this specification. (Updated on 25-
CC MAR-2003 to correct PA field.)
XX
XX Sequence 3010 AA;
SQ

Query Match 86.5%; Score 283; DB 2; Length 3010;
Best Local Similarity 85.7%; Pred. No. 1.4e-27;
Matches 54; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 1 ALENLVNMAASAGTGILMFLVFCAMVYKRLVGATYSLLGLMPLLLLLALPQR 60
DB 747 ALENLVNMAASAGTGILMFLVFCAMVYKRLVGATYSLLGLMPLLLLLALPQR 806
OY 61 AYA 63
DB 807 AYA 809

RESULT 5
AAY06423
ID AAY06423 standard; protein; 3010 AA.
XX
XX AAY06423;
AC
XX
DT 20-MAR-2003 (revised)
DT 27-SEP-1999 (first entry)
XX
XX Non-A, non-B hepatitis virus polypeptide.
DE
XX Non-A, non-B hepatitis B virus; NANBV; antigen; infection; diagnosis;
KW vaccine.
KW
XX
XX Non-A.
OS non-B hepatitis virus.
OS
XX
XX
FH Key Location/Qualifiers
FT 1. .115
FT /note= "core protein"
FT 116. .191
FT /note= "matrix protein"
FT 192. .389
FT /note= "envelope protein"
XX
XX

```

FT Protein 390..729
FT /note="NS1 protein"
FT Protein 730..1006
FT /note="NS2 protein"
FT Protein 1007..1615
FT /note="NS3 protein"
FT Protein 1616..1862
FT /note="NS4a protein"
FT Protein 1863..2013
FT /note="NS4b protein"
FT Protein 2014..3010
FT /note="NS5 protein"
XX
XX EP933426-A1.
XX
XX PD 04-AUG-1999.
XX
XX PF 28-DEC-1990; 99EP-00106005.
XX
XX PR 25-JUN-1990; 90JP-00167466.
XX PR 31-AUG-1990; 90JP-00230921.
XX PR 09-NOV-1990; 90JP-00305605.
XX PR 28-DEC-1990; 90EP-00314371.
XX
XX (OSAU ) UNITV OSAKA.
XX
XX PI Okayama H, Fuke I, Mori C, Takamizawa A, Yoshida I;
XX DR WPI; 1999-407152/35.
XX DR N-PSDB; AAX59394.
XX
XX PT New hepatitis virus polypeptides, useful for diagnosing and treating
XX PT hepatitis infections.
XX
XX PS Claim 2; Fig 2(1) - (16); 56pp; English.
XX
CC This sequence represents the non-A, non-B hepatitis virus (NANBV)
CC polypeptide, as predicted from cDNA (see AAX59394) containing the entire
CC open reading frame of the NANBV genome. To obtain this cDNA, NANBV RNAs
CC were extracted directly from NANBV particles contained in whole blood of
CC a patient having NANBV hepatitis, or from a resected liver of a patient
CC having NANBV hepatitis and liver cancer. The RNA was then converted to
CC double-stranded cDNA. A cDNA library was produced and screened using
CC serum from a convalescent patient having acute NANBV hepatitis and serum
CC from a patient having chronic NANBV hepatitis. The isolated cDNA allows
CC recombinant production of NANBV antigen polypeptides in microbial or
CC eukaryotic cell culture. The method provides the safe production of NANBV
CC antigens with high purity on a large scale at low cost without the
CC biohazard associated with multiplying virus in animals. Claimed NANBV
CC nucleotide sequences are useful for the recombinant production of
CC polypeptides useful as antigens for vaccines, and as diagnostic reagents.
CC (Updated on 20-MAR-2003 to correct PF field.) (Updated on 20-MAR-2003 to
CC correct PR field.)
XX
XX SQ Sequence 3010 AA;
XX
Query Match 86.5%; Score 283; DB 2; Length 3010;
Best Local Similarity 85.7%; Pred. No. 1.4e-27;
Matches 54; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

```

```

AC ADJ82984;
XX
XX DT 06-MAY-2004 (first entry)
XX
XX DE Adenovector construct HCV C-E1-E2-P7-NS2 protein sequence.
XX
XX KW adenovirus construct; vector; hepatitis C virus; viral particle; HCV;
XX KW HCV infection.
XX
XX OS Synthetic.
XX
XX PN WO2004005524-A1.
XX
XX PD 15-JAN-2004.
XX
XX PF 27-JUN-2003; 2003WO-US020409.
XX
XX PR 02-JUL-2002; 2002US-0391167P.
XX
XX PA (MERI ) MERCK & CO INC.
XX
XX PI Ludmerer SW, Flores OA, Graham D, Wolanski BS;
XX DR WPI: 2004-091376/09.
XX DR N-PSDB; ADJ82983.
XX
XX PT Producing hepatitis C virus (HCV) particles comprises incubating Vero
XX PT cells containing a recombinant nucleic acid that comprises an expression
XX PT cassette encoding for at least a HCV C-E1-E2-P7-NS2 sequence.
XX
XX PS Disclosure: Page 37-39; 53pp; English.
XX
CC The present invention relates to a method of producing hepatitis C virus
CC (HCV) particles by incubating Vero cells containing a recombinant nucleic
CC acid that comprises an expression cassette encoding for at least a HCV C-
CC E1-E2-P7-NS2 sequence under conditions suitable for producing the HCV
CC particles. The method is useful for making hepatitis C virus (HCV)
CC particles and measuring the ability of a compound to inhibit HCV particle
CC formation or function. The HCV particles are useful as source materials
CC for obtaining HCV antibodies recognizing a native particle form, and for
CC evaluating the ability of a compound to inhibit particle function or
CC infectivity. The present sequence is a protein encoded by an adenovector
CC construct designated HCV C-E1-E2-P7-NS2 used to demonstrate the method of
CC the invention.
XX
XX SQ Sequence 1026 AA;
XX
Query Match 85.9%; Score 281; DB 8; Length 1026;
Best Local Similarity 85.7%; Pred. No. 7.7e-28;
Matches 54; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

```

```

OY 1 ALENLVNNAASAGTHGILMFLVFCAMVYKGRVPGATYSLLGLMPLLLALLPQR 60
DB 747 ALENLVNNAASAGTHGILMFLVFCAMVYKGRVPGATYSLLGLMPLLLALLPQR 806
OY 61 AYA 63
DB 807 AYA 809

```

```

RESULT 7
AAAY70064
ID AAAY70064 standard; protein; 2307 AA.
XX
XX AC AAAY70064;
XX
XX DT 12-SEP-2003 (revised)
XX DT 05-JUN-2000 (first entry)
XX
XX DE Recombinant fusion pHCAP-1 polypeptide.
XX
XX KW Recombinant plasmid; pHCAP; Hepatitis C virus; HCV; reporter gene;
XX KW NS3 protease; inhibitor; recombinant viral vector; RVV; HCV infection;
XX KW

```



```

FT      Domain                               /label= NS4A/NS4B_cleavage site
FT      1083..1257
FT      /label= NS4B_domain
FT      /note= "Hepatitis C virus non-structural domain"
FT      Cleavage-site                       1259..1278
FT      /label= NS5A/NS5B_cleavage site
FT      Protein                             1282..1784
FT      /label= Secreted alkaline phosphatase protein
FT      /note= "Secreted human placental SNAP"
FT      Region                             1785..1936
FT      /label= pHCAP-1_polypeptide_fragment_3
FT      Region                             1937..2021
FT      /label= pHCAP-1_polypeptide_fragment_4
FT      Region                             2022..2307
FT      /label= pHCAP-1_polypeptide_fragment_5
FT      WO200008469-A1.
XX      PD
XX      17-FEB-2000.
XX      PF
XX      02-AUG-1999; 99WO-US017440.
XX      PR
XX      05-AUG-1998; 98US-00129611.
XX      PR
XX      08-MAR-1999; 99US-00263933.
XX      XX
XX      (AGOU-) AGOURON PHARM INC.
XX      PI
XX      Potts KE, Jackson RL, Patrick AK;
XX      MPI; 2000-224057/19.
XX      DR
XX      N-PSDB; AA251003.
XX      PT
XX      Assessing compounds which augment or inhibit Hepatitis C virus NS3
XX      protease, useful particularly for identifying inhibitors which can be
XX      used for treating Hepatitis C virus infections.
XX      PS
XX      Claim 40; Page 101-108; 153pp; English.
XX      CC
XX      The patent discloses a reporter gene system for use in a cell-based
XX      assessment of inhibitors of Hepatitis C virus (HCV) NS3 protease, using a
XX      recombinant viral vector (RVV), that expresses a secreted human placental
XX      alkaline phosphatase (SEAP) reporter gene polypeptide, under the control
XX      of bacteriophage T7 promoter. The viral vector has been engineered to
XX      express a polypeptide, that includes NS3 HCV serine protease and the
XX      human SEAP gene. This assay system is useful for in vitro screening of
XX      potential protease inhibitors useful in the treatment of HCV infections
XX      and used to evaluate potent NS3 inhibitors, by monitoring the effect of
XX      increasing drug concentration on SEAP activity. NS3 inhibition is
XX      detected as a decrease in SEAP activity. The present sequence is the
XX      recombinant fusion protein, encoded by the RVV pHCAP-3, comprising five
XX      segments and the active NS2 and mutant NS3 protease, fused with the SEAP
XX      reporter protein. The NS3 protease is inactivated by site directed
XX      mutagenesis. The plasmid is constructed using the pTM3 vector and has
XX      been used to generate recombinant vaccinia viruses. (Updated on 12-SEP-
XX      2003 to standardise OS field)
XX      SQ
XX      Sequence 2307 AA;
XX
XX      Query Match      85.9%; Score 281; DB 3; Length 2307;
XX      Best Local Similarity 85.7%; Pred. No. 1,9e-27;
XX      Matches 54; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
XX
XX      QY      1 ALENLVLAASAAGHTGILWLPFCAMVYKGRVLPATYSLGLWPLLLLLALPPR 60
XX      |||||
XX      DB      118 ALENLVLAASAAGHTGILWLPFCAMVYKGRVLPATYSLGLWPLLLLLALPPR 177
XX      |||||
XX      QY      61 AYA 63
XX      |||
XX      DB      178 AYA 180
XX
XX      RESULT 9
XX      AAY70066

```

```

ID      AAY70066 standard; protein; 2307 AA.
XX      AC
XX      AAY70066;
XX      XX
XX      12-SEP-2003 (revised)
XX      DT      05-JUN-2000 (first entry)
XX      XX
XX      DE      Recombinant fusion pHCAP-4 polypeptide.
XX      XX
XX      Recombinant plasmid; pHCAP; Hepatitis C virus; HCV; reporter gene;
XX      KW      NS3 protease; inhibitor; recombinant viral vector; RVV; HCV infection;
XX      KW      secreted alkaline phosphatase; SEAP; serine protease; treatment;
XX      KW      recombinant vaccinia virus.
XX      OS
XX      Hepatitis C virus.
XX      OS
XX      Vaccinia virus.
XX      OS
XX      Enterobacteria phage T7.
XX      OS
XX      Homo sapiens.
XX      OS
XX      Chimeric.
XX      XX
XX      Key
XX      FH      Location/Qualifiers
XX      FT      1..92
XX      FT      /label= pHCAP-1_polypeptide_fragment_1
XX      FT      93..1784
XX      FT      /label= pHCAP-1_polypeptide_fragment_2
XX      FT      93..98
XX      FT      /note= "Additional residues resulting from subcloning of
XX      FT      HCV/SEAP fragment"
XX      FT      100..390
XX      FT      /label= E2/NS2 domain
XX      FT      /note= "Hepatitis C virus non-structural domain"
XX      FT      Misc-difference
XX      FT      364
XX      FT      /note= "Wild type catalytic Cys replaced with Ala, to
XX      FT      inactivate NS2 protease"
XX      FT      391..1028
XX      FT      /label= NS3 domain
XX      FT      /note= "Hepatitis C virus non-structural domain
XX      FT      containing serine protease and helicase enzymes"
XX      FT      536
XX      FT      /note= "Wild type catalytic Ser replaced with Ala, to
XX      FT      inactivate NS3 protease"
XX      FT      1019..1038
XX      FT      /label= NS3/NS4A_cleavage site
XX      FT      1029..1082
XX      FT      /label= NS4A_domain
XX      FT      /note= "Hepatitis C virus non-structural domain"
XX      FT      1076..1092
XX      FT      /label= NS4A/NS4B_cleavage site
XX      FT      1083..1257
XX      FT      /label= NS4B_domain
XX      FT      /note= "Hepatitis C virus non-structural domain"
XX      FT      1259..1278
XX      FT      /label= NS5A/NS5B_cleavage site
XX      FT      1282..1784
XX      FT      /label= Secreted alkaline phosphatase protein
XX      FT      /note= "Secreted human placental SEAP"
XX      FT      Region
XX      FT      1785..1936
XX      FT      /label= pHCAP-1_polypeptide_fragment_3
XX      FT      1937..2021
XX      FT      /label= pHCAP-1_polypeptide_fragment_4
XX      FT      2022..2307
XX      FT      /label= pHCAP-1_polypeptide_fragment_5
XX      PN
XX      WO200008469-A1.
XX      PD
XX      17-FEB-2000.
XX      PF
XX      02-AUG-1999; 99WO-US017440.
XX      PR
XX      05-AUG-1998; 98US-00129611.
XX      PR
XX      08-MAR-1999; 99US-00263933.
XX      PA
XX      (AGOU-) AGOURON PHARM INC.

```

XX Potte KE, Jackson RL, Patick AK;
 XX
 XX MPI: 2000-224057/19.
 DR N-PSDB; AA251004.
 XX
 PT Assessing compounds which augment or inhibit Hepatitis C virus NS3
 PT protease, useful particularly for identifying inhibitors which can be
 PT used for treating Hepatitis C virus infections.
 XX
 PS Claim 41; Page 132-139; 153pp; English.
 XX
 CC The patent discloses a reporter gene system for use in a cell-based
 CC assessment of inhibitors of Hepatitis C virus (HCV) NS3 protease, using a
 CC recombinant viral vector (RVV), that expresses a secreted human placental
 CC alkaline phosphatase (SEAP) reporter gene polypeptide, under the control
 CC of bacteriophage T7 promoter. The viral vector has been engineered to
 CC express a polypeptide, that includes NS3 HCV serine protease and the
 CC human SEAP gene. This assay system is useful for in vitro screening of
 CC potential protease inhibitors useful in the treatment of HCV infections
 CC and used to evaluate potent NS3 inhibitors, by monitoring the effect of
 CC increasing drug concentration on SEAP activity. NS3 inhibition is
 CC detected as a decrease in SEAP activity. The present sequence is the
 CC recombinant fusion protein, encoded by the RVV pHCAP-4, comprising five
 CC segments and the mutant inactive NS2 and mutant inactive NS3 protease
 CC polypeptides, fused with the SEAP reporter protein. The NS2-NS3 domain
 CC are inactivated by site directed mutagenesis. The plasmid is constructed,
 CC using the pTM vector and has been used to generate recombinant vaccinia
 CC viruses. (Updated on 12-SEP-2003 to standardise OS field)
 XX
 SQ Sequence 2307 AA;
 XX
 Query Match 85.9%; Score 281; DB 3; Length 2307;
 Best Local Similarity 85.7%; Pred. No. 1.9e-27;
 Matches 54; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
 OY 1 ALENLVVINAASAAAGTGTGILMFLVFFCAAWYVKGRLVPGATYSLLGLMPLLLLLALPPR 60
 DB 118 ALENLVVINAASVAGAHGILSFLVFFCAAWYIKGRLVPGAAVYALGVWPLLLLLALPPR 177
 OY 61 AYA 63
 DB 178 AYA 180
 XX
 RESULT 10
 ABG32458
 ID ABG32458 standard; protein; 3010 AA.
 XX
 AC ABG32458;
 XX
 DT 15-NOV-2002 (first entry)
 XX
 DE Hepatitis C virus Con 1 isolate polypeptide mutant #7.
 XX
 KM HCV, Con 1; adaptive mutation; liver failure; cirrhosis; mutant; mutcin;
 KM hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
 KM internal ribosome entry site; IRES; NS5A; HCV replication; polypeptide.
 OS Hepatitis C virus.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 2197 /note="Wild-type Phe substituted by Ser"
 XX
 PN WO200259321-A2.
 XX
 PD 01-AUG-2002.
 XX
 PF 16-JAN-2002; 2002WO-BP000526.
 XX
 PR 23-JAN-2001; 2001US-0263479P.

XX
 XX (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.
 XX
 XX De Francesco R, Migliaccio G, Paonessa G;
 PI
 XX MPI: 2002-599793/64.
 DR
 XX
 PT New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
 PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
 PT ribosome entry site (IRES) region, useful in studying HCV replication and
 PT expression.
 XX
 PS Claim 1; Page; 69pp; English.
 XX
 CC The invention relates to nucleic acid molecules comprising altered HCV
 CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
 CC internal ribosome entry site (IRES) region coding for one or more NS3,
 CC NS5A, or EMCV IRES mutations, respectively. The location of the mutations
 CC are detailed in the specification. Also included are (1) an expression
 CC vector comprising a nucleotide sequence coding for the altered nucleic
 CC acids, which is transcriptionally coupled to an exogenous promoter; (2) a
 CC recombinant cell human hepatoma cell comprising the altered nucleic acids
 CC ; (3) a recombinant cell produced by introducing into a human hepatoma
 CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
 CC replicon enhanced cell or which containing a functional HCV replicon; (5)
 CC an HCV replicon enhanced cells made in the method; and (6) measuring the
 CC ability of a compound to affect HCV activity. The HCV replicons and HCV
 CC replicon enhanced cells are useful in studying HCV replication and
 CC expression, and HCV and host cell interactions, producing HCV RNA and
 CC proteins, and providing a system for measuring the ability of a compound
 CC to modulate one or more HCV activities e.g. to discover drugs which may
 CC treat HCV mediated diseases such as liver failure, cirrhosis and
 CC hepatocellular carcinoma. The present sequence is the HCV replicon Con 1
 CC polypeptide (comprising the Core, E1, E2, p7, NS2, NS3, NS4A, NS4B, NS5A
 CC and NS5B proteins), NS5A mutant of the invention. Note: The present
 CC sequence is not shown in the specification but was created by the indexer
 CC using the HCV sequence appearing as ABG32451 and the information in claim
 CC 1
 XX
 SQ Sequence 3010 AA;
 XX
 Query Match 85.9%; Score 281; DB 5; Length 3010;
 Best Local Similarity 85.7%; Pred. No. 2.6e-27;
 Matches 54; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
 OY 1 ALENLVVINAASAAAGTGTGILMFLVFFCAAWYVKGRLVPGATYSLLGLMPLLLLLALPPR 60
 DB 747 ALENLVVINAASVAGAHGILSFLVFFCAAWYIKGRLVPGAAVYALGVWPLLLLLALPPR 806
 OY 61 AYA 63
 DB 807 AYA 809
 XX
 RESULT 11
 ABG32459
 ID ABG32459 standard; protein; 3010 AA.
 XX
 AC ABG32459;
 XX
 DT 15-NOV-2002 (first entry)
 XX
 DE Hepatitis C virus Con 1 isolate polypeptide mutant #8.
 XX
 KM HCV, Con 1; adaptive mutation; liver failure; cirrhosis; mutant; mutcin;
 KM hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
 KM internal ribosome entry site; IRES; NS5A; HCV replication; polypeptide.
 OS Hepatitis C virus.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 2198


```

XX Hepatitis C virus Con 1 isolate polyprotein mutant #4.
XX
XX HCV, Con 1; adaptive mutation; liver failure; cirrhosis; mutant; mutein;
KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
XX Internal ribosome entry site; IRES; NS5A; HCV replication; polyprotein.
XX
OS Hepatitis C virus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 2041
FT /note= "Wild-type Asn substituted by Thr"
XX
XX WO200259321-A2.
XX
XX 01-AUG-2002.
XX
XX 16-JAN-2002; 2002WO-EP000526.
XX
XX 23-JAN-2001; 2001US-0263479P.
XX
XX (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
XX
XX De Francesco R, Migliaccio G, Paonessa G;
XX
XX WPI, 2002-599793/64.
XX
XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
PT ribosome entry site (IRES) region, useful in studying HCV replication and
PT expression.
XX
XX Claim 1; Page: 69pp; English.
XX
XX The invention relates to nucleic acid molecules comprising altered HCV
CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
CC internal ribosome entry site (IRES) region coding for one or more NS3,
CC NS5A, or EMCV IRES mutations, respectively. The location of the mutations
CC are detailed in the specifications. Also included are (1) an expression
CC vector comprising a nucleotide sequence coding for the altered nucleic
CC acids, which is transcriptionally coupled to an exogenous promoter; (2) a
CC recombinant cell human hepatoma cell comprising the altered nucleic acids
CC ; (3) a recombinant cell produced by introducing into a human hepatoma
CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
CC replicon enhanced cell or which containing a functional HCV replicon; (5)
CC an HCV replicon enhanced cells made in the method; and (6) measuring the
CC ability of a compound to affect HCV activity. The HCV replicons and HCV
CC replicon enhanced cells are useful in studying HCV replication and
CC expression, and HCV and host cell interactions, producing HCV RNA and
CC proteins, and providing a system for measuring the ability of a compound
CC to modulate one or more HCV activities e.g. to discover drugs which may
CC treat HCV mediated diseases such as liver failure, cirrhosis and
CC hepatocellular carcinoma. The present sequence is the HCV replicon Con 1
CC polyprotein (comprising the Core, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A
CC and NS5B proteins), NS5A mutant of the invention. Note: The present
CC sequence is not shown in the specification but was created by the indexer
CC using the HCV sequence appearing as ABG32451 and the information in claim
CC 1
XX
XX Sequence 3010 AA:
XX
Query Match 85.9%; Score 281; DB 5; Length 3010;
Best Local Similarity 85.7%; Pred. No. 2.6e-27;
Matches 54; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
OY 1 ALENIIVNAAASAGTHGILWIFECAMWYKGLVFGATSLGLPMLLLALPOR 60
DB 747 ALENIIVNAAASVAGAHGILFLVFCAMWYKGLVFGATSLGLPMLLLALPOR 806
OY 61 AYA 63
DB 807 AYA 809

```

```

RESULT 14
ABG32457
ID ABG32457 standard; protein; 3010 AA.
XX
XX ABG32457;
AC
XX
DT 15-NOV-2002 (first entry)
XX
XX Hepatitis C virus Con 1 isolate polyprotein mutant #6.
DB
XX
XX HCV, Con 1; adaptive mutation; liver failure; cirrhosis; mutant; mutein;
KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
XX Internal ribosome entry site; IRES; NS5A; HCV replication; polyprotein.
XX
OS Hepatitis C virus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 2173
FT /note= "Wild-type Phe substituted by Ser"
XX
XX WO200259321-A2.
XX
XX 01-AUG-2002.
XX
XX 16-JAN-2002; 2002WO-EP000526.
XX
XX 23-JAN-2001; 2001US-0263479P.
XX
XX (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
XX
XX De Francesco R, Migliaccio G, Paonessa G;
XX
XX WPI, 2002-599793/64.
XX
XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
PT ribosome entry site (IRES) region, useful in studying HCV replication and
PT expression.
XX
XX Claim 1; Page: 69pp; English.
XX
XX The invention relates to nucleic acid molecules comprising altered HCV
CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
CC internal ribosome entry site (IRES) region coding for one or more NS3,
CC NS5A, or EMCV IRES mutations, respectively. The location of the mutations
CC are detailed in the specifications. Also included are (1) an expression
CC vector comprising a nucleotide sequence coding for the altered nucleic
CC acids, which is transcriptionally coupled to an exogenous promoter; (2) a
CC recombinant cell human hepatoma cell comprising the altered nucleic acids
CC ; (3) a recombinant cell produced by introducing into a human hepatoma
CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
CC replicon enhanced cell or which containing a functional HCV replicon; (5)
CC an HCV replicon enhanced cells made in the method; and (6) measuring the
CC ability of a compound to affect HCV activity. The HCV replicons and HCV
CC replicon enhanced cells are useful in studying HCV replication and
CC expression, and HCV and host cell interactions, producing HCV RNA and
CC proteins, and providing a system for measuring the ability of a compound
CC to modulate one or more HCV activities e.g. to discover drugs which may
CC treat HCV mediated diseases such as liver failure, cirrhosis and
CC hepatocellular carcinoma. The present sequence is the HCV replicon Con 1
CC polyprotein (comprising the Core, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A
CC and NS5B proteins), NS5A mutant of the invention. Note: The present
CC sequence is not shown in the specification but was created by the indexer
CC using the HCV sequence appearing as ABG32451 and the information in claim
CC 1
XX
XX Sequence 3010 AA:
XX
Query Match 85.9%; Score 281; DB 5; Length 3010;
Best Local Similarity 85.7%; Pred. No. 2.6e-27;

```

Oy		Matches	54;	Conservative	3;	Mismatches	6,	Indels	0;	Gaps	
Dd		Db	747	ALENIUVINAAASACTHGIWMVLVFPFCAMYYKGLVPENTSLGMPLLLLLALPRL		:	:	:	:	:	: 806
Oy			61	AYA 63		:	:	:	:	:	:
						:	:	:	:	:	:
Dd			807	AYA 809		:	:	:	:	:	:
 RESULT_15 ID ABG32460 standard; protein; 3010 AA.											
AC			ABG32460;								
DT			15-NOV-2002	(first entry)							
DE			Hepatitis C virus Con 1 isolate polypotein mutant #9.								
KW			HCV; Con 1; adaptive mutation; liver failure; cirrhosis; mutant; mutein;								
XW			hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;								
KM			internal ribosome entry site; IRES; NSSA; HCV replication; polyprotein.								
OS			Hepatitis C virus.								
SX			Synthetic.								
FY			Key	Location/Qualifiers							
FT			Misc-difference	2199 /note= "Wild-type Ala substituted by Thr"							
NN			WO200259321-A2.								
PD			01-AUG-2002.								
PB			16-JAN-2002; 2002MO-BP000526.								
PR			23-JAN-2001; 2001US-0263479P.								
PA			(RICE-) IST RICERCHE BIOL MOLECOLAIRE ANGELETTI.								
PI			De Francesco R, Migliaccio G, Paonessa G;								
DR			WFI; 2002-599793/64.								
PT			New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV								
PT			NS5 encoding region, or encephalomyocarditis virus (EMCV) internal								
PT			ribosome entry site (IRES) region, useful in studying HCV replication and								
PT			expression.								
PS			Claim 1; Page; 69pp; English.								
XX			The invention relates to nucleic acid molecules comprising altered HCV								
CC			NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV).								
CC			Internal ribosome entry site (IRES) region coding for one or more NS3,								
CC			NSSA, or EMCV IRES mutations, respectively. The location of the mutations								
CC			are detailed in the specification. Also included are (1) an expression								
CC			vector comprising a nucleotide sequence coding for the altered nucleic								
CC			acids, which is transcriptionally coupled to an exogenous promoter; (2) a								
CC			recombinant cell human hepatoma cell comprising the altered nucleic acids								
CC			; (3) a recombinant cell produced by introducing into a human hepatoma								
CC			cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)								
CC			replicon enhanced cell or which containing a functional HCV replicon; (5)								
CC			an HCV replicon enhanced cells made in the method; and (6) measuring the								
CC			ability of a compound to affect HCV activity. The HCV replicons and HCV								
CC			replicon enhanced cells are useful in studying HCV replication and								
CC			expression, and HCV and host cell interactions, producing HCV RNA and								
CC			proteins, and providing a system for measuring the ability of a compound								
CC			to modulate one or more HCV activities e.g. to discover drugs which may								
CC			treat HCV mediated diseases such as liver failure, cirrhosis and								
CC			hepatocellular carcinoma. The present sequence is the HCV replicon Con 1								
CC			polyprotein (comprising the Core, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A								

```
CC      and NS5B proteins), NS5A mutant of the invention. Note: The present  
CC      sequence is not shown in the specification but was created by the indexer  
CC      using the HCV sequence appearing as AB032451 and the information in claim  
CX      1  
CX      xx  
SQ      Sequence 3010 AA;  
  
Query Match          85.9%; Score 281; DB 5; Length 3010;  
Best Local Similarity 85.7%; Pred. No. 2.6e-27;  
Matches   54; Conservative    3; Mismatches   6; Indels    0; Gaps     0  
  
QY      1 ALENLVVILNNAASAAGTHGIIIMFLVFPCFCAWYKGLVPGATYSLTGLMPLLLILLALPOR 60  
        |||||  
DB      747 ALENTLVVILNNAASVAGAHGITLSFLVFPCFCAWYIKGRSLVPGAAVALYGVPMLLLLLLALPPR 806  
        |||||  
QY      61 AYA 63  
        |||  
DB      807 AYA 809
```

Search completed: October 4, 2005, 18:56:23

Job time : 168 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OW protein - protein search, using BW model

Run on: October 4, 2005, 18:52:30 ; Search time 44 Seconds

(without alignments)
106.884 Million cell updates/sec

Title: US-10-669-175-1

Perfect score: 327

Sequence: 1 ALENLVVYNASAGTHGIL.....LLGLWPLLLLLALPQRAYA 63

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents: AA: *
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep: *
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep: *
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep: *
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep: *
5: /cgn2_6/ptodata/1/1aa/6C.COMB.pep: *
6: /cgn2_6/ptodata/1/1aa/6D.COMB.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	283	86.5	88	3	US-09-263-933-23
2	283	86.5	88	4	US-09-919-901-23
3	283	86.5	88	4	US-10-191-966-23
4	283	86.5	277	1	US-08-324-977-38
5	283	86.5	277	2	US-08-384-616-38
6	283	86.5	277	2	US-08-904-686A-38
7	283	86.5	277	3	US-09-315-850-38
8	283	86.5	480	1	US-08-440-542-22
9	283	86.5	480	1	US-08-440-542-22
10	283	86.5	480	1	US-08-231-368-22
11	283	86.5	480	1	US-08-440-210-22
12	283	86.5	480	3	US-09-046-604-22
13	283	86.5	2013	2	US-08-324-977-12
14	283	86.5	2013	2	US-08-384-616-12
15	283	86.5	2013	2	US-08-904-686A-12
16	283	86.5	2013	3	US-09-315-850-12
17	283	86.5	2620	2	US-08-324-977-32
18	283	86.5	2620	2	US-08-384-616-32
19	283	86.5	2620	2	US-08-904-686A-32
20	283	86.5	2620	3	US-09-315-850-32
21	283	86.5	2621	1	US-08-324-977-36
22	283	86.5	2621	2	US-08-384-616-36
23	283	86.5	2621	2	US-08-904-686A-36
24	283	86.5	2621	3	US-09-315-850-36
25	283	86.5	3010	1	US-08-324-977-2
26	283	86.5	3010	1	US-08-324-977-14
27	283	86.5	3010	2	US-08-384-616-2

28	283	86.5	3010	2	US-08-384-616-14	Sequence 14, Appl
29	283	86.5	3010	2	US-08-904-686A-2	Sequence 2, Appl
30	283	86.5	3010	2	US-08-904-686A-14	Sequence 14, Appl
31	283	86.5	3010	3	US-09-315-850-2	Sequence 2, Appl
32	283	86.5	3010	3	US-09-315-850-14	Sequence 14, Appl
33	281	85.9	1692	3	US-09-263-933-4	Sequence 4, Appl
34	281	85.9	1692	3	US-09-263-933-11	Sequence 11, Appl
35	281	85.9	1692	3	US-09-263-933-18	Sequence 18, Appl
36	281	85.9	1692	4	US-09-919-901-4	Sequence 4, Appl
37	281	85.9	1692	4	US-09-919-901-11	Sequence 11, Appl
38	281	85.9	1692	4	US-09-919-901-18	Sequence 18, Appl
39	281	85.9	1692	4	US-10-191-966-4	Sequence 4, Appl
40	281	85.9	1692	4	US-10-191-966-11	Sequence 11, Appl
41	281	85.9	1692	4	US-10-191-966-18	Sequence 18, Appl
42	281	85.9	2307	3	US-09-263-933-2	Sequence 2, Appl
43	281	85.9	2307	3	US-09-263-933-9	Sequence 9, Appl
44	281	85.9	2307	3	US-09-263-933-16	Sequence 16, Appl
45	281	85.9	2307	4	US-09-919-901-2	Sequence 2, Appl

ALIGNMENTS

```
RESULT 1
US-09-263-933-23
Sequence 23, Application US/09263933
Patent No. 6280340
GENERAL INFORMATION:
APPLICANT: Potts, Karen E.
APPLICANT: Jackson, Roberta L.
TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
TITLE OF INVENTION: OF INHIBITORS OF THE HEPATITIS C VIRUS PROTEASE
FILE REFERENCE: 0125-0005A
CURRENT APPLICATION NUMBER: US/09/263, 933
CURRENT FILING DATE: 1999-03-08
EARLIER APPLICATION NUMBER: 09/129, 611
EARLIER FILING DATE: 1998-08-05
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 23
LENGTH: 88
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: amino acid
OTHER INFORMATION: Fragment of the HCV polypeptide
US-09-263-933-23
Query Match 86.5%; Score 283; DB 3; Length 88;
Best Local Similarity 85.7%; Pred. No. 6.2e-29;
Matches 54; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
QY 1 ALENLVVYNASAGTHGILWLVFPCAWYVKGRLVGATYSLGLWPLLLLLALPQRAYA 60
DB 19 ALENLVVYNASVAGAGHILSFVLFPCAWYIKKRLVGATYALYGVWPLLLLLALPQRAYA 78
QY 61 AYA 63
DB 79 AYA 81
RESULT 2
US-09-919-901-23
Sequence 23, Application US/09919901
Patent No. 6599738
GENERAL INFORMATION:
APPLICANT: Potts, Karen E.
APPLICANT: Jackson, Roberta L.
APPLICANT: Patick, Amy K.
TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
TITLE OF INVENTION: OF INHIBITORS OF THE HEPATITIS C VIRUS PROTEASE
FILE REFERENCE: 0125-0005A
```

```
/ CURRENT APPLICATION NUMBER: US/09/919,901
/ CURRENT FILING DATE: 2001-08-02
/ PRIOR APPLICATION NUMBER: 09/263,933
/ PRIOR FILING DATE: 1999-02-08
/ PRIOR APPLICATION NUMBER: 09/129,611
/ PRIOR FILING DATE: 1998-08-05
/ NUMBER OF SEQ ID NOS: 33
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO: 23
/ LENGTH: 88
/ TYPE: PR
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: amino acid
/ OTHER INFORMATION: Fragment of the HCV polyprotein
US-09-919-901-23
```

```
Query Match      86.5%; Score 283; DB 4; Length 88;
Best Local Similarity 85.7%; Pred. No. 6.2e-29;
Matches 54; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 ALENLVNNAASAGTHGILWFLVFCAMVYKGRVPGATYSLLGLWPLLLLLALPOR 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 19 ALENLVNNAASAGTHGILWFLVFCAMVYKGRVPGATYALGVWPLLLLLALPOR 78
    |||||
QY 61 AYA 63
    |||
DB 79 AYA 81
```

```
RESULT 3
US-10-191-966-23
/ Sequence 23, Application US/10191966
/ Patent No. 6790612
/ GENERAL INFORMATION:
/ APPLICANT: Potts, Karen E.
/ APPLICANT: Jackson, Roberta L.
/ APPLICANT: Patrick, Amy K.
/ TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
/ TITLE OF INVENTION: OF INHIBITORS OF THE HEPATITIS C VIRUS PROTEASE
/ FILE REFERENCE: 0125-0005A
/ CURRENT APPLICATION NUMBER: US/10/191,966
/ CURRENT FILING DATE: 2002-07-10
/ PRIOR APPLICATION NUMBER: US/09/263,933
/ PRIOR FILING DATE: 1999-03-08
/ PRIOR APPLICATION NUMBER: 09/129,611
/ PRIOR FILING DATE: 1998-08-05
/ NUMBER OF SEQ ID NOS: 33
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO: 23
/ LENGTH: 88
/ TYPE: PR
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: amino acid
/ OTHER INFORMATION: Fragment of the HCV polyprotein
US-10-191-966-23
```

```
Query Match      86.5%; Score 283; DB 4; Length 88;
Best Local Similarity 85.7%; Pred. No. 6.2e-29;
Matches 54; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 ALENLVNNAASAGTHGILWFLVFCAMVYKGRVPGATYSLLGLWPLLLLLALPOR 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 19 ALENLVNNAASAGTHGILWFLVFCAMVYKGRVPGATYALGVWPLLLLLALPOR 78
    |||||
QY 61 AYA 63
    |||
DB 79 AYA 81
```

```
RESULT 4
US-08-324-977-38
```

```
/ Sequence 38, Application US/08324977
/ Patent No. 5747319
/ GENERAL INFORMATION:
/ APPLICANT: OKAYAMA, Hiroto
/ APPLICANT: FUKE, Isao
/ APPLICANT: MORI, Chisato
/ APPLICANT: TAKAMIZAWA, Akahisa
/ APPLICANT: YOSHIDA, Iwao
/ TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
/ TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
/ NUMBER OF SEQUENCES: 50
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Armstrong, Westerman, Hattori, Mclelland &
/ ADDRESSEE: Naughton
/ STREET: 1725 K St., N.W. Suite 1000
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: U.S.A.
/ ZIP: 20006
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
/ SOFTWARE: ASCII
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/324,977
/ FILING DATE: 18-OCT-1994
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: JP 2-167466
/ FILING DATE: 25-JUN-1990
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: JP 2-230921
/ FILING DATE: 31-AUG-1990
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: JP 2-305605
/ FILING DATE: 09-NOV-1990
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/099,706
/ FILING DATE: 30-JUL-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/769,996
/ FILING DATE: 02-OCT-1991
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/635,451
/ FILING DATE: 28-DEC-1990
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Stevens-Smith, Theresa M.
/ REGISTRATION NUMBER: 36,281
/ REFERENCE/DOCKET NUMBER: 900703D
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 659-2930
/ TELEFAX: (202) 887-0357
/ TELEX: 440142
/ INFORMATION FOR SEQ ID NO: 38:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 277 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-08-324-977-38
```

```
Query Match      86.5%; Score 283; DB 1; Length 277;
Best Local Similarity 85.7%; Pred. No. 2.4e-28;
Matches 54; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 ALENLVNNAASAGTHGILWFLVFCAMVYKGRVPGATYSLLGLWPLLLLLALPOR 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 18 ALENLVNNAASAGTHGILWFLVFCAMVYKGRVPGATYALGVWPLLLLLALPOR 77
    |||||
QY 61 AYA 63
    |||
DB 78 AYA 80
```


Qy 61 AYA 63
|||
Db 78 AYA 80

RESULT 7

US-09-315-850-38
; Sequence 38, Application US/09315850
; Patent No. 6217872
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKE, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, Mclelland &
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44mb
; OPERATING SYSTEM: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/315,850
; FILING DATE:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US/08/904,686
; FILING DATE: 01-AUG-1997
; APPLICATION NUMBER: US 08/324,977
; FILING DATE: 18-OCT-1994
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Mclelland, Le-Nhung
; REGISTRATION NUMBER: 31,541
; REFERENCE/DOCKET NUMBER: 900703G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 277 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-315-850-38

Query Match 86.5%; Score 283; DB 3; Length 277;
Best Local Similarity 85.7%; Pred. No. 2.4e-28;
Matches 54; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ALENLVINLPAASAGTHGLWFLVFCAMWYKGRLVGATYSLGLWPLLLLLALPQR 60
|||
Db 18 ALENLVINLSASVAGAHGLSLVFCAMWIKGRLVGATYALGVWPLLLLLALPQR 77

Qy 61 AYA 63
|||
Db 78 AYA 80

RESULT 8

US-08-440-103-22
; Sequence 22, Application US/08440103
; Patent No. 5670152
; GENERAL INFORMATION:
; APPLICANT: Weiner, Amy J.
; APPLICANT: Houghton, Michael
; TITLE OF INVENTION: Immunoreactive Polypeptide Compositions
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,103
; FILING DATE: 12-MAY-1995
; CLASSIFICATION: 424
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US/08/231,368
; FILING DATE:
; APPLICATION NUMBER: 13-SEP-1991
; FILING DATE: 13-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0205.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 480 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-440-103-22

Query Match 86.5%; Score 283; DB 1; Length 480;
Best Local Similarity 85.7%; Pred. No. 4.6e-28;
Matches 54; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ALENLVINLPAASAGTHGLWFLVFCAMWYKGRLVGATYSLGLWPLLLLLALPQR 60
|||
Db 378 ALENLVINLSASVAGAHGLSLVFCAMWIKGRLVGATYALGVWPLLLLLALPQR 437

Qy 61 AYA 63
|||
Db 438 AYA 440

RESULT 9

US-08-440-542-22
; Sequence 22, Application US/08440542
; Patent No. 5670153
; GENERAL INFORMATION:
; APPLICANT: Weiner, Amy J.
; APPLICANT: Houghton, Michael
; TITLE OF INVENTION: Immunoreactive Polypeptide Compositions
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,542
; FILING DATE: 12-MAY-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/231,368
; FILING DATE:
; APPLICATION NUMBER: US 07/759,575
; FILING DATE: 13-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0205.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 480 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-440-542-22

Query Match 86.5%; Score 283; DB 1; Length 480;
Best Local Similarity 85.7%; Pred. No. 4.6e-28;
Matches 54; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 ALENLVNANASACTGIIWLFVFCAMVYKGRIVPGATYSILGLWPLLLLLALPQR 60
DB 378 ALENLVNANASVAGAGHILSLVFCAMWIKGRIVPGATYALYGVWPLLLLLALPQR 437

QY 61 AYA 63
DB 438 AYA 440

RESULT 10
US-08-231-368-22
; Sequence 22, Application US/08231368
; Patent No. 5756312
; GENERAL INFORMATION:
; APPLICANT: Weiner, Amy J.
; APPLICANT: Houghton, Michael
; TITLE OF INVENTION: Immunoreactive Polypeptide Compositions
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA

ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/231,368
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/759,575
; FILING DATE: 13-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0205.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 480 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-231-368-22

Query Match 86.5%; Score 283; DB 1; Length 480;
Best Local Similarity 85.7%; Pred. No. 4.6e-28;
Matches 54; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 ALENLVNANASACTGIIWLFVFCAMVYKGRIVPGATYSILGLWPLLLLLALPQR 60
DB 378 ALENLVNANASVAGAGHILSLVFCAMWIKGRIVPGATYALYGVWPLLLLLALPQR 437

QY 61 AYA 63
DB 438 AYA 440

RESULT 11
US-08-440-210-22
; Sequence 22, Application US/08440210
; Patent No. 576845
; GENERAL INFORMATION:
; APPLICANT: Weiner, Amy J.
; APPLICANT: Houghton, Michael
; TITLE OF INVENTION: Immunoreactive Polypeptide Compositions
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,210
; FILING DATE: 12-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/231,368
; FILING DATE:
; APPLICATION NUMBER: US 07/759,575
; FILING DATE: 13-SEP-1991
; ATTORNEY/AGENT INFORMATION:

NAME: McClung, Barbara G.
REGISTRATION NUMBER: 33,113
REFERENCE/DOCKET NUMBER: 0205.001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2708
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ. ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 480 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-440-210-22

Query Match 86.5%; Score 283; DB 1; Length 480;
Best Local Similarity 85.7%; Pred. No. 4.6e-28;
Matches 54; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 ALENLVNMAASAGTHGLMFLVFCAMVYKGRVPGATYSLLGLMPLLLLLALPQR 60
DB 378 ALENLVNMAASAGTHGLMFLVFCAMVYKGRVPGATYSLLGLMPLLLLLALPQR 437

QY 61 AYA 63
DB 438 AYA 440

RESULT 12
US-09-046-604-22
Sequence 22, Application US/09046604
Patent No. 6303292
GENERAL INFORMATION:
APPLICANT: Weiner, Amy J.
APPLICANT: Houghton, Michael
TITLE OF INVENTION: Immunoreactive polypeptide compositions
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/046,604
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/231,368
FILING DATE:
APPLICATION NUMBER: US 07/759,575
FILING DATE: 13-SEP-1991
ATTORNEY/AGENT INFORMATION:
NAME: McClung, Barbara G.
REGISTRATION NUMBER: 33,113
REFERENCE/DOCKET NUMBER: 0205.001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2708
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ. ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 480 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-046-604-22

Query Match 86.5%; Score 283; DB 3; Length 480;
Best Local Similarity 85.7%; Pred. No. 4.6e-28;
Matches 54; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 ALENLVNMAASAGTHGLMFLVFCAMVYKGRVPGATYSLLGLMPLLLLLALPQR 60
DB 378 ALENLVNMAASAGTHGLMFLVFCAMVYKGRVPGATYSLLGLMPLLLLLALPQR 437

QY 61 AYA 63
DB 438 AYA 440

RESULT 13
US-08-324-977-12
Sequence 12, Application US/08324977
Patent No. 5747339
GENERAL INFORMATION:
APPLICANT: OKAYAMA, Hiroto
APPLICANT: FUKU, Isao
APPLICANT: MORI, Chisato
APPLICANT: TAKAMIZAWA, Akahisa
APPLICANT: YOSHIDA, Iwao
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Armstrong, Westerman, Hattori, Mclelland &
STREET: 1725 K St. N.W. Suite 1000
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/324,977
FILING DATE: 18-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-230921
FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/099,706
FILING DATE: 30-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Stevens-Smith, Theresa M.
REGISTRATION NUMBER: 36,281
REFERENCE/DOCKET NUMBER: 900703D
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
TELEX: 440142
INFORMATION FOR SEQ. ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2013 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-324-977-12

Query Match 86.5%; Score 283; DB 1; Length 2013;
Best Local Similarity 85.7%; Pred. No. 2.5e-27;
Matches 54; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 1 ALENLVYNASAGTHGILFLVFFCAAYVKGLVPGATYSILGLPRLILLALPFR 60
DB 747 ALENLVYNASVAGAHGILSLVFFCAAWYIKGLVPGATYALGVWPLILLALPFR 806

OY 61 AYA 63
DB 807 AYA 809

RESULT 14
US-08-384-616-12
Sequence 12, Application US/08384616
Patent No. 58477101
GENERAL INFORMATION:
APPLICANT: OKAYAMA, Hiroto
APPLICANT: FUKU, Isao
APPLICANT: MORI, Chisato
APPLICANT: TAKAMIZAWA, Akehisa
APPLICANT: YOSHIDA, Iwao
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Armstrong, Westerman, Hattori, Mcleland &
STREET: 1725 K St. N.W. Suite 1000
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/384,616
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-230921
FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Stevens-Smith, Theresa M.
REGISTRATION NUMBER: 36,281
REFERENCE/DOCKET NUMBER: 900703B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
TELEX: 440142
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2013 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-384-616-12

Query Match 86.5%; Score 283; DB 2; Length 2013;
Best Local Similarity 85.7%; Pred. No. 2.5e-27;
Matches 54; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 1 ALENLVYNASAGTHGILFLVFFCAAYVKGLVPGATYSILGLPRLILLALPFR 60
DB 747 ALENLVYNASVAGAHGILSLVFFCAAWYIKGLVPGATYALGVWPLILLALPFR 806

OY 61 AYA 63
DB 807 AYA 809

RESULT 15
US-08-904-686A-12
Sequence 12, Application US/08904686A
Patent No. 598130
GENERAL INFORMATION:
APPLICANT: OKAYAMA, Hiroto
APPLICANT: FUKU, Isao
APPLICANT: MORI, Chisato
APPLICANT: TAKAMIZAWA, Akehisa
APPLICANT: YOSHIDA, Iwao
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Armstrong, Westerman, Hattori, Mcleland &
STREET: 1725 K St. N.W. Suite 1000
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/904,686A
FILING DATE: 01-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/324,977
FILING DATE: 18-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-230921
FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/099,706
FILING DATE: 30-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Mcleland, Je-Nhung
REGISTRATION NUMBER: 31,541
REFERENCE/DOCKET NUMBER: 900703G
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2930

; TELEFAX: (202) 887-0357
 ; INFORMATION FOR SEQ ID NO: 12:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2013 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-904-686A-12

Query Match 86.5%; Score 283; DB 2; Length 2013;
 Best Local Similarity 85.7%; Pred. No. 2.5e-27;
 Matches 54; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy	1	ALENLVNINASAGTHGILWFVPCAAWYKGRIVGATYSLGLMPILLALLPQR	60
		: : : : : : :	
Db	747	ALENLVNINASAGTHGILWFVPCAAWYKGRIVGATYSLGLMPILLALLPQR	806
		: : : : : : :	
Qy	61	AYA 63	
Db	807	AYA 809	

Search completed: October 4, 2005, 18:58:45
 Job time : 47 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: October 4, 2005, 18:53:50 ; Search time 165 Seconds
(without alignments)
158,499 Million cell updates/sec

Title: US-10-669-175-1

Perfect score: 327
Sequence: 1 ALENLVVLMNAAASAGTHGIL.....LLGLWPLLILLALPORAYA 63

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1846076 seqs, 41516000 residues

Total number of hits satisfying chosen parameters: 1846076

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10F_PUBCOMB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US10G_PUBCOMB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	327	100.0	63	US-10-669-175-1	Sequence 1, Appli
2	283	86.5	88	US-09-919-901-23	Sequence 23, Appli
3	283	86.5	88	US-10-191-966-23	Sequence 23, Appli
4	281	85.9	1692	US-09-919-901-4	Sequence 4, Appli
5	281	85.9	1692	US-09-919-901-11	Sequence 11, Appli
6	281	85.9	1692	US-09-919-901-18	Sequence 18, Appli
7	281	85.9	1692	US-10-191-966-4	Sequence 4, Appli
8	281	85.9	1692	US-10-191-966-11	Sequence 11, Appli
9	281	85.9	1692	US-10-191-966-18	Sequence 18, Appli
10	281	85.9	2307	US-09-919-901-2	Sequence 2, Appli
11	281	85.9	2307	US-09-919-901-9	Sequence 9, Appli

12	281	85.9	2307	10	US-09-919-901-16	Sequence 16, Appli
13	281	85.9	2307	14	US-10-191-966-2	Sequence 2, Appli
14	281	85.9	2307	14	US-10-191-966-9	Sequence 9, Appli
15	281	85.9	2307	14	US-10-191-966-16	Sequence 16, Appli
16	281	85.9	3010	15	US-10-467-000-1	Sequence 1, Appli
17	278	85.0	3010	16	US-10-333-449A-34	Sequence 34, Appli
18	272	83.2	315	16	US-10-669-175-4	Sequence 4, Appli
19	272	83.2	315	9	US-09-929-955-5	Sequence 5, Appli
20	272	83.2	315	13	US-10-104-966-5	Sequence 5, Appli
21	272	83.2	315	15	US-10-719-619-5	Sequence 5, Appli
22	272	83.2	315	16	US-10-817-591-5	Sequence 5, Appli
23	272	83.2	3011	9	US-09-929-955-1	Sequence 9, Appli
24	272	83.2	3011	9	US-09-929-955-1	Sequence 9, Appli
25	272	83.2	3011	9	US-09-747-419-20	Sequence 20, Appli
26	272	83.2	3011	13	US-10-104-966-1	Sequence 1, Appli
27	272	83.2	3011	14	US-10-259-775-20	Sequence 20, Appli
28	272	83.2	3011	15	US-10-189-359-14	Sequence 14, Appli
29	272	83.2	3011	15	US-10-226-734-406	Sequence 406, App
30	272	83.2	3011	15	US-10-719-619-1	Sequence 1, Appli
31	272	83.2	3011	16	US-10-817-591-1	Sequence 1, Appli
32	272	83.2	3011	20	US-11-006-313-20	Sequence 20, Appli
33	269	82.3	463	9	US-09-973-025-46	Sequence 46, Appli
34	269	82.3	463	10	US-09-989-303-36	Sequence 46, Appli
35	269	82.3	463	10	US-09-985-808-46	Sequence 46, Appli
36	269	82.3	463	10	US-09-985-808-46	Sequence 46, Appli
37	269	82.3	463	10	US-09-985-808-46	Sequence 46, Appli
38	269	82.3	463	16	US-10-321-798-46	Sequence 46, Appli
39	269	82.3	490	9	US-09-973-025-36	Sequence 36, Appli
40	269	82.3	490	10	US-09-899-303-36	Sequence 36, Appli
41	269	82.3	490	10	US-09-985-808-36	Sequence 36, Appli
42	269	82.3	490	10	US-09-985-808-36	Sequence 36, Appli
43	269	82.3	490	10	US-09-985-808-36	Sequence 36, Appli
44	269	82.3	490	16	US-10-321-798-36	Sequence 36, Appli
45	269	82.3	692	9	US-09-973-025-48	Sequence 48, Appli

ALIGNMENTS

RESULT 1
US-10-669-175-1
; Sequence 1, Application US/10669175
; Publication No. US20040110795A1
; GENERAL INFORMATION:
; APPLICANT: ZITZMAN, NICOLE
; TITLE OF INVENTION: USE OF IMINOSUGAR DERIVATIVES TO INHIBIT ION CHANNEL
; FILE REFERENCE: 080618-0304
; CURRENT APPLICATION NUMBER: US/10/669,175
; PRIOR FILING DATE: 2003-09-23
; PRIOR APPLICATION NUMBER: 60/412,560
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 1
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: consensus amino acid sequence for HCV p7
US-10-669-175-1

Query Match 100.0%; Score 327, DB 16; Length 63;
Best Local Similarity 100.0%; Pred. No. 4, 7e-31;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALENLVVLMNAAASAGTHGILWLFVFCAMVYKGLVGCATYSLGLWPLLILLALPOR 60
DB 1 ALENLVVLMNAAASAGTHGILWLFVFCAMVYKGLVGCATYSLGLWPLLILLALPOR 60
QY 61 AYA 63

Db 61 AYA 63

RESULT 2

US-09-919-901-23
; Sequence 23, Application US/09919901
; Publication No. US20030082518A1
; GENERAL INFORMATION:
; APPLICANT: Potes, Karen E.
; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patrick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; TITLE OF INVENTION: OF INHIBITORS OF THE HEPATITIS C VIRUS PROTEASE
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/09/919,901
; CURRENT FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 09/263,933
; PRIOR FILING DATE: 1999-02-08
; PRIOR APPLICATION NUMBER: 09/129,611
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: amino acid
US-09-919-901-23

Query Match 86.5%; Score 283; DB 10; Length 88;
Best Local Similarity 85.7%; Pred. No. 1.1e-25;
Matches 54; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 ALENLVNAAASAGTHGILWLFVFCAMVYKGRIVGATYSLLGLPPLLLLLALPQR 60
DB 19 ALENLVNAAASAGTHGILWLFVFCAMVYKGRIVGATYSLLGLPPLLLLLALPQR 78
QY 61 AYA 63
DB 79 AYA 81

RESULT 3

US-10-191-966-23
; Sequence 23, Application US/10191966
; Publication No. US20030175692A1
; GENERAL INFORMATION:
; APPLICANT: Potes, Karen E.
; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patrick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; TITLE OF INVENTION: OF INHIBITORS OF THE HEPATITIS C VIRUS PROTEASE
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/10/191,966
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US/09/263,933
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/129,611
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: amino acid
US-10-191-966-23

Query Match 86.5%; Score 283; DB 14; Length 88;
Best Local Similarity 85.7%; Pred. No. 1.1e-25;
Matches 54; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 ALENLVNAAASAGTHGILWLFVFCAMVYKGRIVGATYSLLGLPPLLLLLALPQR 60
DB 19 ALENLVNAAASAGTHGILWLFVFCAMVYKGRIVGATYSLLGLPPLLLLLALPQR 78
QY 61 AYA 63
DB 79 AYA 81

RESULT 4

US-09-919-901-4
; Sequence 4, Application US/09919901
; Publication No. US20030082518A1
; GENERAL INFORMATION:
; APPLICANT: Potes, Karen E.
; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patrick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; TITLE OF INVENTION: OF INHIBITORS OF THE HEPATITIS C VIRUS PROTEASE
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/09/919,901
; CURRENT FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 09/263,933
; PRIOR FILING DATE: 1999-02-08
; PRIOR APPLICATION NUMBER: 09/129,611
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1692
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: :
US-09-919-901-4

Query Match 85.9%; Score 281; DB 10; Length 1692;
Best Local Similarity 85.7%; Pred. No. 3.7e-24;
Matches 54; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ALENLVNAAASAGTHGILWLFVFCAMVYKGRIVGATYSLLGLPPLLLLLALPQR 60
DB 26 ALENLVNAAASAGTHGILWLFVFCAMVYKGRIVGATYSLLGLPPLLLLLALPQR 85
QY 61 AYA 63
DB 86 AYA 88

RESULT 5

US-09-919-901-11
; Sequence 11, Application US/09919901
; Publication No. US20030082518A1
; GENERAL INFORMATION:
; APPLICANT: Potes, Karen E.
; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patrick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; TITLE OF INVENTION: OF INHIBITORS OF THE HEPATITIS C VIRUS PROTEASE
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/09/919,901
; CURRENT FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 09/263,933
; PRIOR FILING DATE: 1999-02-08
; PRIOR APPLICATION NUMBER: 09/129,611
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11

LENGTH: 1692
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: :
US-09-919-901-11

Query Match 85.9%; Score 281; DB 10; Length 1692;
Best Local Similarity 85.7%; Pred. No. 3.7e-24;
Matches 54; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ALENLVVLMNAAAGTGILMFLVFCAMVYKGRLVGATYSLLGLMPLLLLLALPQR 60
DB 26 ALENLVVLMNAAAGTGILMFLVFCAMVYKGRLVGATYSLLGLMPLLLLLALPQR 85
QY 61 AYA 63
DB 86 AYA 88

RESULT 6

US-09-919-901-18
Sequence 18, Application US/09919901
Publication No. US20030082518A1
GENERAL INFORMATION:
APPLICANT: Potts, Karen E.
APPLICANT: Jackson, Roberta L.
APPLICANT: Patick, Amy K.
TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
FILE REFERENCE: 0125-0005A
CURRENT APPLICATION NUMBER: US/09/919,901
CURRENT FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: 09/263,933
PRIOR FILING DATE: 1999-02-08
PRIOR APPLICATION NUMBER: 09/129,611
PRIOR FILING DATE: 1998-08-05
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 18
LENGTH: 1692
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: :
US-09-919-901-18

Query Match 85.9%; Score 281; DB 10; Length 1692;
Best Local Similarity 85.7%; Pred. No. 3.7e-24;
Matches 54; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ALENLVVLMNAAAGTGILMFLVFCAMVYKGRLVGATYSLLGLMPLLLLLALPQR 60
DB 26 ALENLVVLMNAAAGTGILMFLVFCAMVYKGRLVGATYSLLGLMPLLLLLALPQR 85
QY 61 AYA 63
DB 86 AYA 88

RESULT 7

US-10-191-966-4
Sequence 4, Application US/10191966
Publication No. US20030175692A1
GENERAL INFORMATION:
APPLICANT: Potts, Karen E.
APPLICANT: Jackson, Roberta L.
APPLICANT: Patick, Amy K.
TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
FILE REFERENCE: 0125-0005A
CURRENT APPLICATION NUMBER: US/10/191,966
CURRENT FILING DATE: 2002-07-10

PRIOR APPLICATION NUMBER: US/09/263,933
PRIOR FILING DATE: 1999-03-08
PRIOR APPLICATION NUMBER: 09/129,611
PRIOR FILING DATE: 1998-08-05
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 1692
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: :
US-10-191-966-4

Query Match 85.9%; Score 281; DB 14; Length 1692;
Best Local Similarity 85.7%; Pred. No. 3.7e-24;
Matches 54; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ALENLVVLMNAAAGTGILMFLVFCAMVYKGRLVGATYSLLGLMPLLLLLALPQR 60
DB 26 ALENLVVLMNAAAGTGILMFLVFCAMVYKGRLVGATYSLLGLMPLLLLLALPQR 85
QY 61 AYA 63
DB 86 AYA 88

RESULT 8

US-10-191-966-11
Sequence 11, Application US/10191966
Publication No. US20030175692A1
GENERAL INFORMATION:
APPLICANT: Potts, Karen E.
APPLICANT: Jackson, Roberta L.
APPLICANT: Patick, Amy K.
TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
FILE REFERENCE: 0125-0005A
CURRENT APPLICATION NUMBER: US/10/191,966
CURRENT FILING DATE: 2002-07-10
PRIOR APPLICATION NUMBER: 09/263,933
PRIOR FILING DATE: 1999-03-08
PRIOR APPLICATION NUMBER: 09/129,611
PRIOR FILING DATE: 1998-08-05
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 11
LENGTH: 1692
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: :
US-10-191-966-11

Query Match 85.9%; Score 281; DB 14; Length 1692;
Best Local Similarity 85.7%; Pred. No. 3.7e-24;
Matches 54; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ALENLVVLMNAAAGTGILMFLVFCAMVYKGRLVGATYSLLGLMPLLLLLALPQR 60
DB 26 ALENLVVLMNAAAGTGILMFLVFCAMVYKGRLVGATYSLLGLMPLLLLLALPQR 85
QY 61 AYA 63
DB 86 AYA 88

RESULT 9

US-10-191-966-18
Sequence 18, Application US/10191966
Publication No. US20030175692A1
GENERAL INFORMATION:
APPLICANT: Potts, Karen E.

```

; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; TITLE OF INVENTION: OF INHIBITORS OF THE HEPATITIS C VIRUS PROTEASE
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/10/191,966
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US/09/263,933
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/129,611
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 1692
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION:
US-10-191-966-18
```

```

Query Match      85.9%; Score 281; DB 14; Length 1692;
Best Local Similarity 85.7%; Pred. No. 3.7e-24;
Matches 54; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
```

```
QY      1 ALENLVINAASAGTGILMFLVFPCAMVYKGRIVPGATYSLGLWPLLLLLALPOR 60
        |||||
DB      26 ALENLVINAASVAGAHGILSFVFCAMWIKGRIVGAAVALGVWPLLLLLALPPR 85
        |||||
QY      61 AYA 63
        |||
DB      86 AYA 88
```

```

RESULT 10
US-09-919-901-2
; Sequence 2, Application US/09919901
; Publication No. US20030082518A1
; GENERAL INFORMATION:
; APPLICANT: Potte, Karen E.
; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; TITLE OF INVENTION: OF INHIBITORS OF THE HEPATITIS C VIRUS PROTEASE
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/09/919,901
; CURRENT FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 09/263,933
; PRIOR FILING DATE: 1999-02-08
; PRIOR APPLICATION NUMBER: 09/129,611
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2307
; LENGTH: 2307
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION:
US-09-919-901-2
```

```

Query Match      85.9%; Score 281; DB 10; Length 2307;
Best Local Similarity 85.7%; Pred. No. 5.1e-24;
Matches 54; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
```

```
QY      1 ALENLVINAASAGTGILMFLVFPCAMVYKGRIVPGATYSLGLWPLLLLLALPOR 60
        |||||
DB      118 ALENLVINAASVAGAHGILSFVFCAMWIKGRIVGAAVALGVWPLLLLLALPPR 177
        |||||
QY      61 AYA 63
        |||
DB      178 AYA 180
```

```

RESULT 11
US-09-919-901-9
; Sequence 9, Application US/09919901
; Publication No. US20030082518A1
; GENERAL INFORMATION:
; APPLICANT: Potte, Karen E.
; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; TITLE OF INVENTION: OF INHIBITORS OF THE HEPATITIS C VIRUS PROTEASE
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/09/919,901
; CURRENT FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 09/263,933
; PRIOR FILING DATE: 1999-02-08
; PRIOR APPLICATION NUMBER: 09/129,611
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 2307
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION:
US-09-919-901-9
```

```

Query Match      85.9%; Score 281; DB 10; Length 2307;
Best Local Similarity 85.7%; Pred. No. 5.1e-24;
Matches 54; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
```

```
QY      1 ALENLVINAASAGTGILMFLVFPCAMVYKGRIVPGATYSLGLWPLLLLLALPOR 60
        |||||
DB      118 ALENLVINAASVAGAHGILSFVFCAMWIKGRIVGAAVALGVWPLLLLLALPPR 177
        |||||
QY      61 AYA 63
        |||
DB      178 AYA 180
```

```

RESULT 12
US-09-919-901-16
; Sequence 16, Application US/09919901
; Publication No. US20030082518A1
; GENERAL INFORMATION:
; APPLICANT: Potte, Karen E.
; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; TITLE OF INVENTION: OF INHIBITORS OF THE HEPATITIS C VIRUS PROTEASE
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/09/919,901
; CURRENT FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 09/263,933
; PRIOR FILING DATE: 1999-02-08
; PRIOR APPLICATION NUMBER: 09/129,611
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 2307
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION:
US-09-919-901-16
```

```

Query Match      85.9%; Score 281; DB 10; Length 2307;
Best Local Similarity 85.7%; Pred. No. 5.1e-24;
Matches 54; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
```

```
QY      1 ALENLVINAASAGTGILMFLVFPCAMVYKGRIVPGATYSLGLWPLLLLLALPOR 60
```

Db 118 ALENLVVINAASVAGAHGILSFVFCAMWYIKGRLVGGAAYALYGWVPLLLLLLALPPR 177
QY 61 AYA 63
Db 178 AYA 180

RESULT 13

US-10-191-966-2
; Sequence 2, Application US/10191966
; Publication No. US20030175692A1
; GENERAL INFORMATION:
; APPLICANT: Poter, Karen E.
; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patrick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; TITLE OF INVENTION: OF INHIBITORS OF THE HEPATITIS C VIRUS PROTEASE
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/10/191,966
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US/09/263,933
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/129,611
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2307
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: :
US-10-191-966-2

Query Match 85.9%; Score 281; DB 14; Length 2307;
Best Local Similarity 85.7%; Pred. No. 5.1e-24;
Matches 54; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ALENLVVINAASVAGAHGILSFVFCAMWYIKGRLVGGAAYALYGWVPLLLLLLALPPR 60
Db 118 ALENLVVINAASVAGAHGILSFVFCAMWYIKGRLVGGAAYALYGWVPLLLLLLALPPR 177
QY 61 AYA 63
Db 178 AYA 180

RESULT 14

US-10-191-966-9
; Sequence 9, Application US/10191966
; Publication No. US20030175692A1
; GENERAL INFORMATION:
; APPLICANT: Poter, Karen E.
; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patrick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; TITLE OF INVENTION: OF INHIBITORS OF THE HEPATITIS C VIRUS PROTEASE
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/10/191,966
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US/09/263,933
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/129,611
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 2307
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: :

US-10-191-966-9

Query Match 85.9%; Score 281; DB 14; Length 2307;
Best Local Similarity 85.7%; Pred. No. 5.1e-24;
Matches 54; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ALENLVVINAASVAGAHGILSFVFCAMWYIKGRLVGGAAYALYGWVPLLLLLLALPPR 60
Db 118 ALENLVVINAASVAGAHGILSFVFCAMWYIKGRLVGGAAYALYGWVPLLLLLLALPPR 177
QY 61 AYA 63
Db 178 AYA 180

RESULT 15

US-10-191-966-16
; Sequence 16, Application US/10191966
; Publication No. US20030175692A1
; GENERAL INFORMATION:
; APPLICANT: Poter, Karen E.
; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patrick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; TITLE OF INVENTION: OF INHIBITORS OF THE HEPATITIS C VIRUS PROTEASE
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/10/191,966
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US/09/263,933
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/129,611
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 2307
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: :
US-10-191-966-16

Query Match 85.9%; Score 281; DB 14; Length 2307;
Best Local Similarity 85.7%; Pred. No. 5.1e-24;
Matches 54; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ALENLVVINAASVAGAHGILSFVFCAMWYIKGRLVGGAAYALYGWVPLLLLLLALPPR 60
Db 118 ALENLVVINAASVAGAHGILSFVFCAMWYIKGRLVGGAAYALYGWVPLLLLLLALPPR 177
QY 61 AYA 63
Db 178 AYA 180

Search completed: October 4, 2005, 19:01:35
Job time : 166 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2005, 18:46:50 ; Search time 14 Seconds

(without alignments)
432.975 Million cell updates/sec

Title: US-10-669-175-1

Perfect score: 327

Sequence: 1 ALENLVNANASAACTHGL.....LLGLWPLLLLLALPQRAYA 63

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	283	86.5	3010	1 GNMVTC	genome polypeptide
2	277	84.7	3010	1 S18030	genome polypeptide
3	275	84.1	3010	1 GNMVCH	genome polypeptide
4	272	83.2	3011	1 GNMVCH	genome polypeptide
5	271	82.9	3010	1 A45573	genome polypeptide
6	271	82.9	3010	1 GNMVTC	genome polypeptide
7	265	81.0	3011	1 GNMVTC	genome polypeptide
8	264	80.7	3011	1 S40770	genome polypeptide
9	234	71.6	3033	1 GNMVTC	genome polypeptide
10	227	69.4	3033	1 JQ1303	genome polypeptide
11	212	64.8	3014	1 JCS620	genome polypeptide
12	166	50.8	3014	1 S19876	genome polypeptide
13	164	50.2	782	2 S18032	genome polypeptide
14	161	49.2	782	2 PNO677	hypothetical prote
15	160	48.9	782	2 S19875	genome polypeptide
16	158	48.3	782	2 S18031	genome polypeptide
17	157	48.3	782	2 AC2320	hypothetical prote
18	157	48.3	782	2 S76656	hypothetical prote
19	157	48.3	782	2 T04367	plasma membrane in
20	157	48.3	782	2 S32804	beta-3-adrenergic
21	157	48.3	782	2 A75417	hypothetical prote
22	157	48.3	782	2 S65459	beta-3-adrenergic
23	157	48.3	782	2 S44900	ZK652.10 protein -
24	157	48.3	782	2 S44742	C02C2.4 protein -
25	157	48.3	782	2 AE3024	hypothetical prote
26	157	48.3	782	2 D98260	inner membrane pro
27	157	48.3	782	2 G82069	conserved hypothet
28	157	48.3	782	2 S73116	comet protein - My
29	157	48.3	782	2 GRH08	beta-3-adrenergic

30	60.5	18.5	414	1 ORH083	beta-3-adrenergic
31	60.5	18.5	454	2 AD2460	hypothetical prote
32	60.5	18.5	1337	2 T38949	hypothetical prote
33	60	18.3	285	2 T12342	major intrinsic pr
34	60	18.3	326	2 T30166	hypothetical prote
35	60	18.3	341	2 E96019	probable sugar upt
36	60	18.3	400	2 A41679	beta-3-adrenergic
37	60	18.3	400	2 A53281	hypothetical prote
38	60	18.3	459	2 H83675	hypothetical prote
39	60	18.3	1085	2 H82511	probable hemolysin
40	59.5	18.2	219	2 H75342	conserved hypothet
41	59.5	18.2	283	2 E71310	MFS permease [prol
42	59.5	18.2	516	2 AB3088	proline/betaine tr
43	59.5	18.2	516	2 G98198	probable integral
44	59.5	18.2	665	2 C81439	hypothetical prote
45	59	18.0	133	2 PH0210	

ALIGNMENTS

RESULT 1

GNMVTC
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstruct)
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C:Accession: A38465
R:Takamizawa, A.; Mori, C.; Fuke, I.; Manabe, S.; Murakami, S.; Fujita, J.; Onishi, E.;
J. Virol. 65, 1105-1113, 1991
A>Title: Structure and organization of the hepatitis C virus genome isolated from human
A:Reference number: A38465; MUID:911406598; PMID:1847440
A:Accession: A38465
A:Molecule type: genomic RNA
A:Residues: 1-3010 <TRK>
A:Cross-references: UNIPROT:P26663; EMBL:M58335; NID:G329770; PIDN:AAA72945.1; PID:G3297
C:Superfamily: hepatitis C virus genome polypeptide
C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructura
F:2-15/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted
F:192-389/Product: major envelope protein E #status predicted <ME>
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1007-1615/Product: hepacivirin #status predicted <NS3>
F:1230-1337/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4a #status predicted <NS4>
F:1863-2013/Product: nonstructural protein NS4b #status predicted <NS4>
F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,2077,22

Query Match
Best local similarity 86.5%; Score 283; DB 1; Length 3010;
Matches 54; Conservative 4; Mismatch 5; Indels 0; Gaps 0;

QY 1 ALENLVNANASAACTHGLWPLLLALPQRAYA 60
DB 747 ALENLVNANASAACTHGLWPLLLALPQRAYA 806
QY 61 AYA 63
DB 807 AYA 809

RESULT 2
S18030
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstruct)
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
A:Variety: isolate JX1

C.Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
 C.Accession: S18030, S35370, A48332, S18029
 R.Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
 Submitted to the EMBL Data Library, September 1991
 A.Description: A whole genome of hepatitis C virus cDNA was isolated from a single patient
 A.Reference number: S18028
 A.Accession: S18030
 A.Molecule type: genomic RNA
 A.Residues: 1-3010 <HON>
 A.Cross-references: UNIPROT:Q6949; EMBL:X61596; NID:G59478; PIDN:CAA43793.1; PID:G59479
 A.Experimental source: isolate JX1 from an individual
 R.Honda, M.; Kaneko, S.; Uonuma, M.; Kobayashi, K.; Murakami, S.
 Arch. Virol. 128, 163-169, 1993
 A.Title: Sequence analysis of putative structural regions of hepatitis C virus isolated
 A.Reference number: A48332; MUID:93119270; PMID:8380322
 A.Accession: S35370
 A.Molecule type: genomic RNA
 A.Residues: 1-547, 'T', 549-621, 'V', 623-624, 'S', 626-652, 'DL', 655-761, 'T', 763-782 <HON>
 A.Cross-references: EMBL:X61591
 A>Note: this sequence is inconsistent with the nucleotide translation
 A>Note: the authors translated the codon AGG for residue 43 as Pro, TGG for residue 320 as Trp, and TTC for residue 771 as Ser
 A>Note: sequence extracted from NCBI backbone (NCBIN:121747, NCBI:121748)
 C.Superfamily: hepatitis C virus genome polypeptide
 C.Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polypeptide; serin
 F.2-115/Product: capsid protein C #status predicted <CPC>
 F.116-191/Product: envelope protein M #status predicted <EPM>
 F.192-389/Product: major envelope protein E #status predicted <MEB>
 F.390-729/Product: nonstructural protein NS1 #status predicted <NS1>
 F.730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
 F.1007-1615/Product: nonstructural protein NS3 #status predicted <NS3>
 F.1230-1237/Region: nucleotide-binding motif A (P-loop)
 F.1312-1317/Region: nucleotide-binding motif B
 F.1316-1319/Region: DEXH motif
 F.1616-1862/Product: nonstructural protein NS4 #status predicted <NS4>
 F.1863-2013/Product: nonstructural protein NS5 #status predicted <NS5>
 F.2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
 F.196,209,224,250,305,417,423,448,532,540,556,576,623,645/Binding site: carbohydrate (As

Query Match Best Local Similarity 84.7%; Score 277; DB 1; Length 3010;
 Matches 53; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 ALENTLVNAAAGAGHGLMFLVPCAAWYKGRVPCATYSLGLMPLLLALLPORA 60
 DB 747 ALENTLVNAAAGAGHGLMFLVPCAAWYKGRVPCATYSLGLMPLLLALLPORA 806

QY 61 AYA 63
 DB 807 AYA 809

RESULT 3
 GNMVCT
 genome polypeptide - hepatitis C virus (strain J)
 N:Contains: capsid protein C; envelope protein M; major envelope protein E; nonstructural
 protein NS4; nonstructural protein NS4b; nonstructural protein NS5
 C:Species: hepatitis C virus
 C.Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
 C.Accession: A39253; PS0086
 R.Kato, N.; Hijioka, M.; Ootsubo, Y.; Nakagawa, M.; Ohkoshi, S.; Sugimura, T.; Shinoh
 Proc. Natl. Acad. Sci. U.S.A. 87, 9524-9528, 1990
 A.Title: Molecular cloning of the human hepatitis C virus genome from Japanese patients
 A.Reference number: A39253; MUID:91088550; PMID:2175903
 A.Accession: A39253
 A.Molecule type: genomic RNA
 A.Residues: 1-3010 <KAT>
 A.Cross-references: UNIPROT:P26662; GB:D90208; NID:G221610; PIDN:BA414233.1; PID:G221611
 R.Kato, N.; Ohkoshi, S.; Shimotohno, K.
 Proc. Jpn. Acad. 65B, 219-223, 1989
 A.Title: Japanese isolates of the non-A, non-B hepatitis viral genome show sequence vari
 A.Reference number: PS0085
 A.Accession: PS0086

A.Molecule type: genomic RNA
 A.Residues: 2650-2707 <KA2>
 A.Experimental source: Japanese isolate
 C.Comment: The cleavage sites of this polypeptide have not been determined.
 C.Superfamily: hepatitis C virus genome polypeptide
 C.Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polypeptide; serin
 F.2-115/Product: capsid protein C #status predicted <CPC>
 F.116-191/Product: envelope protein M #status predicted <EPM>
 F.192-389/Product: major envelope protein E #status predicted <MEB>
 F.390-729/Product: nonstructural protein NS1 #status predicted <NS1>
 F.730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
 F.1007-1615/Product: nonstructural protein NS3 #status predicted <NS3>
 F.1230-1237/Region: nucleotide-binding motif A (P-loop)
 F.1312-1317/Region: nucleotide-binding motif B
 F.1316-1319/Region: DEXH motif
 F.1616-1862/Product: nonstructural protein NS4 #status predicted <NS4>
 F.1863-2013/Product: nonstructural protein NS4b #status predicted <NS4b>
 F.2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
 F.196,209,224,250,305,325,417,423,430,448,532,556,576,623,645,1213,1255,2041,2077,2240,2

Query Match Best Local Similarity 84.1%; Score 275; DB 1; Length 3010;
 Matches 52; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 2 ALENTLVNAAAGAGHGLMFLVPCAAWYKGRVPCATYSLGLMPLLLALLPORA 61
 DB 748 ALENTLVNAAAGAGHGLMFLVPCAAWYKGRVPCATYSLGLMPLLLALLPORA 807

QY 62 YA 63
 DB 808 YA 809

RESULT 4
 GNMVCT
 genome polypeptide - hepatitis C virus (strain H)
 N:Contains: capsid protein C; envelope protein M; hepatitis B (EC 3.4.21.98) (nonstructu
 protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
 C:Species: hepatitis C virus
 A>Note: host Homo sapiens (man)
 C.Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
 C.Accession: A36814; A41546
 R.Inchausti, G.; Zebadee, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
 Submitted to GenBank, July 1992
 A.Description: Genomic structure of the human prototype strain H of hepatitis C virus: c
 A.Reference number: A36814
 A.Accession: A36814
 A.Molecule type: genomic RNA
 A.Residues: 1-3011 <INC>
 A.Cross-references: UNIPROT:P27958; GB:M67463; NID:G329737; PIDN:AAA5534.1; PID:G329738
 R.Inchausti, G.; Zebadee, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
 Proc. Natl. Acad. Sci. U.S.A. 89, 10292-10296, 1991
 A.Title: Genomic structure of the human prototype strain H of hepatitis C virus: compari
 A.Reference number: A41546; MUID:92052256; PMID:1658800
 A.Contents: annotation
 A>Note: neither amino acid nor nucleotide sequence is given
 C.Superfamily: hepatitis C virus genome polypeptide
 C.Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructura
 F.1-115/Product: capsid protein C #status predicted <CPC>
 F.116-191/Product: envelope protein M #status predicted <EPM>
 F.192-389/Product: major envelope protein E #status predicted <MEB>
 F.390-729/Product: nonstructural protein NS1 #status predicted <NS1>
 F.730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
 F.1007-1615/Product: nonstructural protein NS3 #status predicted <NS3>
 F.1230-1237/Region: nucleotide-binding motif A (P-loop)
 F.1312-1317/Region: nucleotide-binding motif B
 F.1316-1319/Region: DEXH motif
 F.1616-1862/Product: nonstructural protein NS4 #status predicted <NS4>
 F.1863-2013/Product: nonstructural protein NS4b #status predicted <NS4b>
 F.2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>
 F.196,209,224,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,2240,23

Query Match Best Local Similarity 83.2%; Score 272; DB 1; Length 3011;

A:Molecule type: genomic RNA
A:Residues: 1-3010 <CH>
A:Cross-references: UNIPROT:P29846; GB:M84754
C:Superfamily: hepatitis C virus genome polypeptide
C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructural
F:1-115/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <EPM>
F:192-389/Product: major envelope protein E #status predicted <ME>
F:730-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1007-1615/Product: hepacivirin #status predicted <NS3>
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEHH motif
F:1616-1862/Product: nonstructural protein NS4 #status predicted <NA>
F:1863-2013/Product: nonstructural protein NS4B #status predicted <NS4B>
F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,223,224,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,207

Query Match 82.9%; Score 271; DB 1; Length 3010;
Best Local Similarity 82.5%; Pred. No. 2,1e-22;
Matches 52; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

OY 1 ALENVTVNAAASAGHGILMPLVFPCAAYVVGRTVPGATYSILGLWPLLLLTALPOR 60
Db 747 ALENLVFNAAASVAGMHGTLSFLVFCPCAWYIKGRVLPGAVALYGWPLLLTLPPR 806

OY 61 AYA 63
|||
Db 807 AYA 809

RESULT 7
GNMVJ3
genome polypeptide - hepatitis C virus (strain HCV-1)
N:Contains: capsid protein C; envelope protein M; hepacivirin (BC 3.4.21.98) (nonstructu
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 09-Jul-2004
C:Accession: AJ39166; PQ0403; PQ0404
R:Choo, Q.L.; Richman, K.H.; Han, J.H.; Berger, K.; Lee, C.; Dong, C.; Gallegos, C.; Col
Proc. Natl. Acad. Sci. U.S.A. 88, 2451-2455, 1991
A>Title: Genetic organization and diversity of the hepatitis C virus.
A:Reference number: AJ39166; WUID:91172826; PMID:1848704
A:Accession: AJ39166
A:Molecule type: mRNA
A:Residues: 1-3011 <CH>
A:Cross-references: UNIPROT:P26664; GB:M62321; NID:g329873; PIDN:AAA5676.1; PID:g329874
R:Chan, S.W.; McOmish, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yap, P.L
J. Gen. Virol. 73, 1131-1141, 1992
A>Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship to e
A:Reference number: PQ0393; WUID:92268871; PMID:1316939
A:Accession: PQ0403
A:Molecule type: genomic RNA
A:Residues: 1577-1633 <CHA>
A:Cross-references: DDBJ:D10128
A:Experimental source: Isolates E-B16
A:Accession: PQ0404
A>Status: preliminary
A:Molecule type: genomic RNA
A:Residues: 1577-1633 <CH2>
A:Experimental source: Isolates E-B17
C:Superfamily: hepatitis C virus genome polypeptide
C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructura
F:1-115/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <EPM>
F:192-389/Product: major envelope protein E #status predicted <ME>
F:730-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1007-1615/Product: hepacivirin #status predicted <NS3>
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEHH motif

A:Residues: 1-3033 <OKA>
A:Cross-references: UNIPROT:P26660; GB:D00944; NID:G221650; PIDN:BAA00792.1; PID:G221651
A:Experimental source: isolate HC-J6 from a Japanese individual
C:Superfamily: hepatitis C virus genome polypeptide
C:Keywords: ATP; glycoprotein; hydroxylase; P-loop; polypeptide; serine proteinase; transmembrane
F:2-115/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <EPM>
F:193-389/Product: major envelope protein E #status predicted <MES>
F:390-733/Product: nonstructural protein NS1 #status predicted <NS1>
F:734-1010/Product: nonstructural protein NS2 #status predicted <NS2>
F:1011-1619/Product: hepatitis C virus genome polypeptide <NS3>
F:11316-13321/Product: nucleotide-binding motif B
F:1320-13323/Product: nucleotide-binding motif B
F:1620-1866/Product: nonstructural protein NS4a #status predicted <NS4a>
F:1867-2017/Product: nonstructural protein NS4b #status predicted <NS4b>
F:2018-1033/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,224,305,325,417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,2038,28

Query Match 69.4%; Score 227; DB 1; Length 3033;
Best Local Similarity 71.4%; Pred. No. 2e-17;
Matches 45; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

Qy 1 ALENLVVINAASAGTHGILMFLVPCAAWYKGRVPGATYSLLGLMPLLLILLALPQR 60
Db 751 ALEKLVVHAASASGNGFLVFVIFVAAWYKGRVPLATYSLLGLMPLLLILLALPQR 810

Qy 61 AYA 63
Db 811 AYA 813

RESULT 11
JCS620
genome polypeptide - hepatitis C virus (isolate ETH1480)
N:Contains: capsid protein C; envelope protein M; hepatitis C virus genome polypeptide (EC 3.4.21.98) (nonstructural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C:Accession: JCS620
R:Chamberlain, R.W.; Adams, N.J.; Taylor, L.A.; Simmonds, P.; Elliott, R.M.
Biochem. Biophys. Res. Commun. 236, 44-49, 1997
A:Title: The complete coding sequence of hepatitis C virus genome polypeptide 5a, the predominant
A:Reference number: JCS620; MUID:97365593; PMID:9223423
A:Accession: JCS620
A:Molecule type: mRNA
A:Residues: 1-3014 <CHN>
A:Cross-references: UNIPROT:O39928; GB:Y13184
A:Experimental source: genotype 5a, which predominates in South Africa
A>Note: the translation of the nucleotide sequence is not complete in this paper
C:Superfamily: hepatitis C virus genome polypeptide
C:Keywords: ATP; glycoprotein; hydroxylase; nucleotide binding; P-loop; polypeptide; serine
F:2-115/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <EPM>
F:193-389/Product: major envelope protein E #status predicted <MES>
F:384-408/Product: hypervariable #status predicted
F:390-730/Product: nonstructural protein NS1 #status predicted <NS1>
F:733-1007/Product: nonstructural protein NS2 #status predicted <NS2>
F:1008-1616/Product: hepatitis C virus genome polypeptide <NS3>
F:1121-1238/Product: nucleotide-binding motif A (P-loop)
F:11313-1318/Product: nucleotide-binding motif B
F:11317-1320/Product: nucleotide-binding motif B
F:1617-1863/Product: nonstructural protein NS4a #status predicted <NS4a>
F:1864-2014/Product: nonstructural protein NS4b #status predicted <NS4b>
F:2015-3014/Product: nonstructural protein NS5 #status predicted <NS5>
F:2210-2249/Product: interferon sensitivity determining #status predicted

Query Match 64.8%; Score 212; DB 1; Length 3014;
Best Local Similarity 63.9%; Pred. No. 9.9e-16;
Matches 39; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

Qy 3 ALENLVVINAASAGTHGILMFLVPCAAWYKGRVPGATYSLLGLMPLLLILLALPQR 62
Db 750 ALEKLVVHAASASGNGFLVFVIFVAAWYKGRVPLATYSLLGLMPLLLILLALPQR 809

Qy 63 A 63
Db 810 A 810

RESULT 12
S19876
genome polypeptide - hepatitis C virus (isolate JK5) (fragment)
N:Contains: core protein; envelope protein 1; nonstructural protein 2; NS1/E2 protein
C:Species: hepatitis C virus
A:Variety: isolate JK5
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C:Accession: S19876
R:Honda, M.; Kaneo, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
submitted to the EMBL Data Library, September 1991
A:Description: Sequence analysis of putative structural regions of Hepatitis C Virus iso
A:Reference number: S18029
A:Accession: S19876
A:Molecule type: genomic RNA
A:Residues: 1-782 <HON>
A:Cross-references: UNIPROT:Q68953; EMBL:X61595; NID:G59486; PIDN:CAA43792.1; PID:G59487
A:Experimental source: isolate JK5
C:Superfamily: hepatitis C virus genome polypeptide
C:Keywords: capsid protein; core protein; envelope protein; glycoprotein; nonstructural
F:1-191/Product: core protein #status predicted <MAT1>
F:193-383/Product: envelope protein 1 #status predicted <MAT2>
F:384-733/Product: NS1/E2 protein #status predicted <MAT3>
F:734-782/Product: nonstructural protein 2 (fragment) #status predicted <MAT4>

Query Match 50.8%; Score 166; DB 2; Length 782;
Best Local Similarity 88.9%; Pred. No. 4.6e-11;
Matches 32; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ALENLVVINAASAGTHGILMFLVPCAAWYKGR 36
Db 747 ALENLVVINAASAGTHGILMFLVPCAAWYKGR 782

RESULT 13
S18032
genome polypeptide - hepatitis C virus (isolate JK4) (fragment)
N:Contains: core protein; envelope protein 1; nonstructural protein 2; NS1/E2 protein
C:Species: hepatitis C virus
A:Variety: isolate JK4
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C:Accession: S18032
R:Honda, M.; Kaneo, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
submitted to the EMBL Data Library, September 1991
A:Description: Sequence analysis of putative structural regions of Hepatitis C Virus iso
A:Reference number: S18029
A:Accession: S18032
A:Molecule type: genomic RNA
A:Residues: 1-782 <HON>
A:Cross-references: UNIPROT:Q68952; EMBL:X61594
A:Experimental source: isolate JK4
C:Superfamily: hepatitis C virus genome polypeptide
C:Keywords: capsid protein; core protein; envelope protein; glycoprotein; nonstructural
F:1-191/Product: core protein #status predicted <MAT1>
F:193-383/Product: envelope protein 1 #status predicted <MAT2>
F:384-733/Product: NS1/E2 protein #status predicted <MAT3>
F:734-782/Product: nonstructural protein 2 (fragment) #status predicted <MAT4>

Query Match 50.2%; Score 164; DB 2; Length 782;
Best Local Similarity 86.1%; Pred. No. 7.7e-11;
Matches 31; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ALENLVVINAASAGTHGILMFLVPCAAWYKGR 36
Db 747 ALENLVVINAASAGTHGILMFLVPCAAWYKGR 782

RESULT 14

PNM0677
 hypochlorite protein 787 - hepatitis C virus (fragment)
 C/Species: hepatitis C virus
 C/Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 09-Jul-2004
 C/Accession: PNM0677
 R/Cho, S.H.; Yoon, J.I.; Chang, J.E.; Ahn, B.M.; Lee, C.H.; Lee, Y.I.
 Biochem. Biophys. Res. Commun. 196; 780-788, 1993
 A/Title: Genomic typing of hepatitis C viruses from Korean patients: Implications of gen
 A/Reference number: PNM0677; MUID:94059104; PMID:8240354
 A/Accession: PNM0677
 A/Molecule type: mRNA
 A/Residues: 1-787 <CHO>
 A/Cross-references: UNIPROT:Q0824; GB:L20498; NID:G1381031; PIDN:AA802608.1; PID:G13810
 C/Superfamily: hepatitis C virus genome polypeptide
 C/Keywords: glycoprotein; nonstructural protein
 P:196, 209, 234, 250, 305, 325, 421, 427, 452, 536, 544, 560, 580, 627, 649/Binding site: carbohydrate

Query Match	49.2%	Score 161;	DB 2;	Length 787;
Best Local Similarity	83.8%;	Pred. No. 1.7e-10;		
Matches 31;	Conservative 2;	Mismatches 4;	Indels 0;	Gaps 0;

```
QY      1 ALENLVLNAAASAGTHGILMELVFPCAAMVYKGRIV 37
          ||||| | | ||| ||||| : : |||
Db       751 ALENLVLNAAAYAGAIGILSELFPCAAMVIKGLIV 787
```

RESULT 15

genome polyprotein - hepatitis C virus (isolate JK3) (fragment)
 N/Contains: core protein; envelope protein 1; nonstructural protein 2; NS1/E2 protein
 C/Species: hepatitis C virus
 A/Variety: isolate JK3
 C/Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
 C/Accession: S19875
 R/Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
 submitted to the EMBL Data Library, September 1991
 A/Description: Sequence analysis of putative structural regions of Hepatitis C Virus isolate
 A/Reference number: S18029
 A/Accession: S19875
 A/Molecule type: genomic RNA
 A/Residues: 1-782 <ON>
 A/Cross-references: UNIPROT:068951; EMBL:X61592; NID:G959482; PIDN:CAA43789.1; PID:G959483
 A/Experimental source: isolate JK3
 A/Superfamily: hepatitis C virus genome polyprotein
 C/Keywords: capsid protein; core protein; envelope protein; glycoprotein; nonstructural
 P:1-191/Product: core protein #status predicted <MAT1>
 F:192-383/Product: envelope protein 1 #status predicted <MAT2>
 F:384-733/Product: NS1/E2 protein #status predicted <MAT3>
 F:734-782/Product: nonstructural protein 2 (fragment) #status predicted <MAT4>

Query Match	48.9%	Score 160;	DB 2;	Length 782;
Best Local Similarity	86.1%;	Pred. No. 2.2e-10;		
Matches	31; Conservative	2; Mismatches	3; Indels	0; Gaps 0

```

Qy      1 ALENVLVNAASAGTHGIIMFLVFPCAANYVKGL 36
          |||||  : |||||  : |||||
Db      747 ALENVLVNAASVAGTRGIPFLVFPCAANYIKGL 783

```

Search completed: October 4, 2005, 18:57:53
Job time : 16 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2005, 18:41:15 ; Search time 63 Seconds
(without alignments)
512.079 Million cell updates/sec

Title: US-10-669-175-1
Perfect score: 327
Sequence: 1 ALENLVNNAASAACTHGIL.....LLGLWPLLLLLALPQRAYA 63

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_03:*
1: uniprot_sprot:*
2: uniprot_crembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	286	87.5	3010	2	Q9J3H7 hepatitis c
2	285	87.2	3010	2	Q685J3 hepatitis c
3	284	86.9	3010	2	Q9J3G8 hepatitis c
4	284	86.9	3010	2	Q9J3H1 hepatitis c
5	283	86.5	3010	1	POLG_HCVBK
6	282	86.2	3010	2	Q8V6J8 hepatitis c
7	282	86.2	3010	2	Q9J3F9 hepatitis c
8	282	86.2	3010	2	Q9Q1X3 hepatitis c
9	281	85.9	562	2	Q98UN3 hepatitis c
10	281	85.9	1117	2	Q8Q9P1 hepatitis c
11	281	85.9	3010	2	Q688Z6 hepatitis c
12	281	85.9	3010	2	Q6GYR9 hepatitis c
13	281	85.9	3010	2	Q9DTE1 hepatitis c
14	281	85.9	3010	2	Q9J3G3 hepatitis c
15	281	85.9	3010	2	Q9J3G6 hepatitis c
16	281	85.9	3010	2	Q9Q1X4 hepatitis c
17	281	85.9	3010	2	Q9Q1X5 hepatitis c
18	281	85.9	3010	2	Q9QPM2 hepatitis c
19	281	85.9	3010	2	Q9WMX2 hepatitis c
20	280	85.9	3011	2	Q9DPT8 hepatitis c
21	280	85.6	3010	2	Q81541 hepatitis c
22	280	85.6	3010	2	Q81989 hepatitis c
23	280	85.6	3010	2	Q9J3G0 hepatitis c
24	280	85.6	3010	2	Q9J3G1 hepatitis c
25	280	85.6	3013	2	Q9J3H4 hepatitis c
26	279	85.3	3010	2	Q68285 hepatitis c
27	279	85.3	3010	2	Q68285 hepatitis c
28	279	85.3	3010	2	Q68285 hepatitis c
29	278	85.0	3010	2	Q9J3G2 hepatitis c
30	278	85.0	3010	2	Q9J3G2 hepatitis c
31	278	85.0	3010	2	Q9J3G2 hepatitis c

32	278	85.0	3010	2	Q9J3G5 hepatitis c
33	278	85.0	3013	2	Q9Q1X9 hepatitis c
34	278	85.0	3013	2	Q9Q1Y0 hepatitis c
35	277	84.7	3010	2	P90192 hepatitis c
36	277	84.7	3010	2	P90195 hepatitis c
37	277	84.7	3010	2	Q68949 hepatitis c
38	277	84.7	3010	2	Q81760 hepatitis c
39	277	84.7	3010	2	Q9J3I0 hepatitis c
40	277	84.7	3010	2	Q9Q1X6 hepatitis c
41	277	84.7	3010	2	Q9Q1X7 hepatitis c
42	277	84.7	3010	2	Q9Q1X8 hepatitis c
43	277	84.7	3010	2	Q9Q1Y5 hepatitis c
44	277	84.7	3010	2	Q9Q1Y6 hepatitis c
45	277	84.7	3010	2	Q9WIK8 hepatitis c

ALIGNMENTS

```
RESULT 1
ID Q9J3H7 PRELIMINARY; PRT; 3010 AA.
AC Q9J3H7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage, Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RC STRAIN=MD5;
RA Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;
RL Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF207756; AAF65946.1; -.
DR PIR; A61196; A61196.
DR PIR; P00246; P00246.
DR PIR; P00804; P00804.
DR PIR; PS0329; PS0329.
DR HSP; Q8YJ51; ICWX.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR00345; Cytochrome_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4A.
DR InterPro; IPR001490; HCV NS4B.
DR InterPro; IPR002868; HCV NS5A.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01542; HCV_core; 1.
```

DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV_NSI; 1.
DR Pfam; PF01538; HCV_NSI; 1.
DR Pfam; PF02307; HCV_NSI; 1.
DR Pfam; PF01006; HCV_NSI; 1.
DR Pfam; PF01001; HCV_NSI; 1.
DR Pfam; PF01506; HCV_NSI; 1.
DR Pfam; PF00998; HCV_NSI; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.
KM Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
SQ SEQUENCE 3010 AA; 327366 MW; D863F7317FFA106 CRC64;

Query Match 87.5%; Score 286; DB 2; Length 3010;
Best Local Similarity 87.3%; Pred. No. 1.7e-23;
Matches 55; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 ALENLVNNAASAGTHGILMFLVFCAMVYKGLVPGATYSLGLMPLLLALPDR 60
DB 747 ALENLVNNAASAGTHGILMFLVFCAMVYKGLVPGATYSLGLMPLLLALPDR 806

QY 61 AYA 63
DB 807 AYA 809

RESULT 2
068533 PRELIMINARY; PRT; 3010 AA.
AC 068533;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Mueller H.M., Pfaff E., Heller A.E., Goesser T., Theilmann L.;
RL Submitted (JAN-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL; U5476; AAA6907.1; -
DR PIR; A61196; A61196.
DR HSP; OBYS1; ICWX.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR011492; Flavi_DEAD.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV_NSI.
DR InterPro; IPR000745; HCV_NSI.
DR InterPro; IPR001490; HCV_NSI.
DR InterPro; IPR002868; HCV_NSI.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.

DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NSI.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVlr.
DR Pfam; PF07652; Flavi_DEAD; 1.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV_NSI; 1.
DR Pfam; PF01538; HCV_NSI; 1.
DR Pfam; PF02907; HCV_NSI; 1.
DR Pfam; PF01006; HCV_NSI; 1.
DR Pfam; PF01001; HCV_NSI; 1.
DR Pfam; PF01506; HCV_NSI; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; HCV_NSI; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.
KM Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
SQ SEQUENCE 3010 AA; 327117 MW; 1D0B5A97C1466ED6 CRC64;

Query Match 87.2%; Score 285; DB 2; Length 3010;
Best Local Similarity 87.3%; Pred. No. 2.2e-23;
Matches 55; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 ALENLVNNAASAGTHGILMFLVFCAMVYKGLVPGATYSLGLMPLLLALPDR 60
DB 747 ALENLVNNAASAGTHGILMFLVFCAMVYKGLVPGATYSLGLMPLLLALPDR 806

QY 61 AYA 63
DB 807 AYA 809

RESULT 3
030368 PRELIMINARY; PRT; 3010 AA.
AC 030368;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;
RL Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF207765; AAF65955.1; -
DR PIR; A61196; A61196.
DR HSP; PS0328; PS0329.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV_capsid.

DR InterPro; IPR002521; HCV_core.
 DR InterPro; IPR002519; HCV_env.
 DR InterPro; IPR002531; HCV_NS1.
 DR InterPro; IPR000745; HCV_NS4b.
 DR InterPro; IPR001490; HCV_NS4b.
 DR InterPro; IPR002868; HCV_NS5a.
 DR InterPro; IPR002166; HCV_RdRp.
 DR InterPro; IPR001650; Helicase_C.
 DR InterPro; IPR004109; Peptidase_S29.
 DR InterPro; IPR009003; Pept_Ser_Cys.
 DR InterPro; IPR002518; Pept_U39_HCV_NS2.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR007094; RNA_pol_PSVlr.
 DR Pfam; PF01543; HCV_capsid; 1.
 DR Pfam; PF01542; HCV_core; 1.
 DR Pfam; PF01539; HCV_env; 1.
 DR Pfam; PF01560; HCV_NS1; 1.
 DR Pfam; PF01538; HCV_NS2; 1.
 DR Pfam; PF02907; HCV_NS3; 1.
 DR Pfam; PF01006; HCV_NS4a; 1.
 DR Pfam; PF01001; HCV_NS4b; 1.
 DR Pfam; PF01506; HCV_NS5a; 1.
 DR Pfam; PF00271; Helicase_C; 1.
 DR Pfam; PF00998; Viral_RdRp; 1.
 DR SMART; SM00487; DEXDC; 1.
 DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN_1.
 DR Cost protein; Envelope protein; Glycoprotein; Nonstructural protein; Polyprotein; Transmembrane.
 KM POLYPEPTIDE; 3010 AA; 326863 MW; E0BBB3528215457C CRC64;
 SQ SEQUENCE

Query Match Best Local Similarity 86.9%; Score 284; DB 2; Length 3010;
 Matches 54; Conservative 85.7%; Pred. No. 2.9e-23;
 Mismatches 5; Indels 0; Gaps 0;

OY 1 ALENLVVINAASAGTGILMFLVFCAMVYKGRVPGATYSLGLMPLLLLLALPQR 60
 DB 747 ALENLVVINAASVAGTGILSLFLVFCAMVYKGRVPGATYSLGLMPLLLLLALPQR 806

OY 61 AYA 63
 DB 807 AYA 809

RESULT 4
 ID 09J3H1 PRELIMINARY; PRT; 3010 AA.
 AC 09J3H1;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Polyprotein.
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MD21;
 RA Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.,
 RL Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF207762; AAF65952.1; -.
 DR PIR; A61196; A61196.
 DR PIR; PS0329; PS0329.
 DR HSSP; Q8UTS1; ICMX.
 DR GO; GO:0016021; C:Integral to membrane; IEA.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0008026; F:RNA-dependent helicase activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
 DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.

DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR GO; GO:0019079; P:viral genome replication; IEA.
 DR GO; GO:0019087; P:viral transformation; IEA.
 DR InterPro; IPR000345; CytC_heme_BS.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR011545; DEAD/DEAH_N.
 DR InterPro; IPR002522; HCV_capsid.
 DR InterPro; IPR002521; HCV_core.
 DR InterPro; IPR002519; HCV_env.
 DR InterPro; IPR002531; HCV_NS1.
 DR InterPro; IPR00745; HCV_NS4a.
 DR InterPro; IPR001490; HCV_NS4b.
 DR InterPro; IPR002868; HCV_NS5a.
 DR InterPro; IPR002166; HCV_RdRp.
 DR InterPro; IPR001650; Helicase_C.
 DR InterPro; IPR004109; Peptidase_S29.
 DR InterPro; IPR009003; Pept_Ser_Cys.
 DR InterPro; IPR002518; Pept_U39_HCV_NS2.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR007094; RNA_pol_PSVlr.
 DR Pfam; PF01543; HCV_capsid; 1.
 DR Pfam; PF01542; HCV_core; 1.
 DR Pfam; PF01539; HCV_env; 1.
 DR Pfam; PF01560; HCV_NS1; 1.
 DR Pfam; PF01538; HCV_NS2; 1.
 DR Pfam; PF02907; HCV_NS3; 1.
 DR Pfam; PF01006; HCV_NS4a; 1.
 DR Pfam; PF01001; HCV_NS4b; 1.
 DR Pfam; PF01506; HCV_NS5a; 1.
 DR Pfam; PF00271; Helicase_C; 1.
 DR Pfam; PF00998; Viral_RdRp; 1.
 DR SMART; SM00487; DEXDC; 1.
 DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN_1.
 DR Cost protein; Envelope protein; Glycoprotein; Nonstructural protein; Polyprotein; Transmembrane.
 KM POLYPEPTIDE; 3010 AA; 327119 MW; E48938CA801B97AC CRC64;
 SQ SEQUENCE

Query Match Best Local Similarity 86.9%; Score 284; DB 2; Length 3010;
 Matches 54; Conservative 85.7%; Pred. No. 2.9e-23;
 Mismatches 5; Indels 0; Gaps 0;

OY 1 ALENLVVINAASAGTGILMFLVFCAMVYKGRVPGATYSLGLMPLLLLLALPQR 60
 DB 747 ALENLVVINAASVAGTGILSLFLVFCAMVYKGRVPGATYSLGLMPLLLLLALPQR 806

OY 61 AYA 63
 DB 807 AYA 809

RESULT 5
 ID POLG_HCVBK STANDARD; PRT; 3010 AA.
 AC P26663;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
 DE Envelope glycoprotein E1 (GP35) (GP35); Envelope glycoprotein E2
 DE (GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
 DE (EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirus)
 DE (EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein
 NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein
 DE NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
 OS Hepatitis C virus (isolate BK) (HCV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11105;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91140698; PubMed=1847440;
 RA Takamizawa A., Mori C., Fuke I., Manabe S., Murakami S., Fujita J.,

DR	InterPro: IPR0009003; Pept_Ser_CyB.			
DR	InterPro: IPR002518; Pept_Usg_HCV_NS2.			
DR	InterPro: IPR004109; PeptIdase_S25.			
DR	InterPro: IPR007095; RNA_pol_DS_PS.			
DR	InterPro: IPR007094; RNA_pol_Pevtr.			
DR	Pfam: PF01543; HCV_capsid; 1.			
DR	Pfam: PF01542; HCV_core; 1.			
DR	Pfam: PF01539; HCV_env; 1.			
DR	Pfam: PF01560; HCV_NS1; 1.			
DR	Pfam: PF01563; HCV_NS2; 1.			
DR	Pfam: PF02907; HCV_NS3; 1.			
DR	Pfam: PF01006; HCV_NS4a; 1.			
DR	Pfam: PF01001; HCV_NS4b; 1.			
DR	Pfam: PF01506; HCV_NS5a; 1.			
DR	Pfam: PF00998; Viral_RdRp; 1.			
DR	SMART: SM00487; dExDc; 1.			
KW	3D-structure; ATP-binding; Coat protein; Core protein; Direct protein sequencing; Envelope protein; Glycoprotein; Helicase; Hydroxylase; Nonstructural protein; Polyprotein; RNA-directed RNA polymerase; Serine protease; Transferase; Transmembrane.			
FT	INT_MET	1	1	Removed from capsid protein C by the cellular aminopeptidase.
FT	CHAIN	1	115	Capsid protein C (Potential).
FT	CHAIN	116	191	Matrix protein (Potential).
FT	CHAIN	192	383	Major envelope protein E (Potential).
FT	CHAIN	384	729	Nonstructural protein NS1/E2 (Potential)
FT	CHAIN	730	1006	Nonstructural protein NS2 (Potential).
FT	CHAIN	1007	1615	Protease/helicase NS3 (Potential).
FT	CHAIN	1616	1862	Nonstructural protein NS4a (Potential).
FT	CHAIN	1863	2013	Nonstructural protein NS4b (Potential).
FT	CHAIN	2014	3010	RNA-directed RNA polymerase (Potential).
FT	TRANSMEM	347	369	Potential.
FT	ACT_SITE	1083	1083	Charge relay system.
FT	ACT_SITE	1107	1107	Charge relay system.
FT	ACT_SITE	1165	1165	Charge relay system.
FT	NP_BIND	1230	1237	APP (Potential).
FT	SITE	1316	1319	DECH box.
FT	CARBOHYD	196	196	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	209	209	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	234	234	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	250	250	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	305	305	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	417	417	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	423	423	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	430	430	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	448	448	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	532	532	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	540	540	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	556	556	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	576	576	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	623	623	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	645	645	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	2041	2041	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	2077	2077	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	2240	2240	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	2529	2529	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	2788	2788	N-linked (GlcNAc. . .) (Potential).
FT	STRAND	1031	1035	
FT	HELIX	1039	1047	
FT	STRAND	1050	1050	
FT	STRAND	1059	1063	
FT	STRAND	1068	1074	
FT	TURN	1075	1076	
FT	STRAND	1077	1081	
FT	HELIX	1082	1085	
FT	TURN	1086	1087	
FT	STRAND	1090	1092	
FT	TURN	1093	1094	
FT	STRAND	1095	1097	
FT	STRAND	1101	1103	
FT	TURN	1104	1107	
FT	STRAND	1108	1112	

FT	STRAND	1120	1120
FT	STRAND	1122	1122
FT	STRAND	1129	1133
FT	TURN	1135	1136
FT	STRAND	1139	1144
FT	STRAND	1149	1157
FT	HELIX	1158	1161
FT	TURN	1162	1163
FT	TURN	1165	1166
FT	STRAND	1168	1171
FT	TURN	1172	1174
FT	STRAND	1175	1186
FT	TURN	1187	1188
FT	STRAND	1189	1197
FT	HELIX	1198	1202
FT	TURN	1203	1204
FT	STRAND	1680	1688
FT	STRAND	2421	2421
FT	STRAND	2423	2425
FT	STRAND	2439	2439
FT	HELIX	2444	2449
FT	HELIX	2453	2455
FT	STRAND	2456	2458
FT	HELIX	2461	2463
FT	HELIX	2464	2471
FT	STRAND	2474	2474
FT	HELIX	2481	2494
FT	TURN	2495	2496
FT	STRAND	2498	2498
FT	HELIX	2504	2509
FT	TURN	2510	2510
FT	TURN	2513	2514
FT	TURN	2519	2520
FT	HELIX	2524	2528

```

Query Match Similarity      86.5%: Score 283; DB 1; Length 3010;
Best Local Similarity      85.7%: Pred No. 3.7e-23;
Matches 54; Conservative 4; Mismatches 5; Indels 0; Gaps 0

QY      1 ALENLVVMAASAACTGILWELFVFCAMVYKGRVFGATVSLGLWPLLLLLLPOR 60
Db      747 ALENLVVNSASVAGAGHILSFVFCAMVYKGRVFGATVSLGLWPLLLLLLPOR 806
QY      61 AYA 63
Db      807 AYA 809

RESULT 6
Q8V638
ID      Q8V638      PRELIMINARY;      PRT;      3010 AA.
AC      Q8V638;
DT      01-MAR-2002 (TRENBLrel. 20, Created)
DT      01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT      01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DS      Polypeptidein (Fragment).
OS      Hepatitis C virus.
OC      Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC      Hepacivirus.
NC      NCBI_TextID=11103;
RX      [1]
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Panning L.J., Itakura J., Nagayama K., Enomoto N.;
RL      Submitted (OCT-2000) to the EMBL/Genbank/DBJ databases.
DR      EMBL; AF131916; AAL55821.1; -.
DR      PIR; A61196; A61196.
DR      PIR; P00246; P00246.
DR      PIR; P00804; P00804.
DR      PIR; PS0329; PS0329.
DR      HSSP; OB1YS1; ICWX.
DR      GO; GO:0016021; C:Integral to membrane; IEA.
DR      GO; GO:0019028; C:Viral capsid; IEA.
DR      GO; GO:0019031; C:Viral envelope; IEA.

```

DR	GO: GO:0005524; P:ATP binding; IEA.
DR	GO: GO:0008026; F:ATP-dependent helicase activity; IEA.
DR	GO: GO:0003723; F:RNA binding; IEA.
DR	GO: GO:0003368; F:RNA-directed RNA polymerase activity; IEA.
DR	GO: GO:0008236; F:serine-type peptidase activity; IEA.
DR	GO: GO:0005198; F:structural molecule activity; IEA.
DR	GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
GO	GO: GO:0006350; P:transcription; IEA.
DR	GO: GO:0019079; P:vital genome replication; IEA.
DR	GO: GO:0019087; P:vital transformation; IEA.
DR	Pfam: PF01543; HCV capsid; 1.
DR	Pfam: PF01542; HCV core; 1.
DR	Pfam: PF01539; HCV env; 1.
DR	Pfam: PF01560; HCV NS1; 1.
DR	Pfam: PF01538; HCV NS2; 1.
DR	Pfam: PF02907; HCV NS3; 1.
DR	Pfam: PF01006; HCV NS4; 1.
DR	Pfam: PF01001; HCV NS4b; 1.
DR	Pfam: PF01506; HCV NS5a; 1.
DR	Pfam: PF00271; Helicase_C; 1.
DR	Pfam: PF00998; Viral_RdRP; 1.
DR	SMART: SM00487; DEXDC; 1.
DR	PROSITE: PS00190; CYTOCHROME_C; UNKNOWN 1.
KW	Coat protein; Envelope protein; Glycoprotein; Nonstructural protein.
FT	NON_PTM
EQ	SEQUENCE 3010 3010 33AAAAGC072510839 CEC64;

Query Match	86.2%	Score 282;	DB 2;	Length 3010;
Best Local Similarity	85.7%	Pred. No. 4.8e-23;		
Matches	54;	Conservative	3;	Mismatches 6;
			Indels	0;
			Gaps	0

```

Qy 1 ALENVLVUNAASAGTHGILWFLVFFCAAMVYKRGVLPGATYSLSGLMPILLILLALPOR 60
   |||||
Db 747 ALENVLVUNAASVAGTHGILSLFLVFFCAAMVYKRGVLPGAAVAYVGMVPLILLILLALPPR 806
   |||||

Qy 61 AYA 63
   |||
Db 807 AYA 809

```

```

RESULT 7
0903f9
ID 0903f9 PRELIMINARY; PRT; 3010 AA.
AC 0903f9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polypeptein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_taxid=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MM033;
RA Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.,
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF007774: AAF65964.1; -.
DR PIR: A61196; A61196.
DR PIR: P00246; P00246.
DR PIR: PS0329; PS0329.
DR HSSP: P26664; 1HE1.
DR GO: GO:0016021; C:Integral to membrane; IEA.
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0019031; C:viral envelope; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003688; F:RNA-directed RNA polymerase activity; IEA.
DR GO: GO:0008236; F:serine-type peptidase activity; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.

```

DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR000345; C:ycb_heme_BS.
DR InterPro; IPR001410; DEAD_
DR InterPro; IPR01545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR00745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_NS5a.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_core; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; Viral_RdRp; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
SQ SEQUENCE 3010 AA; 327103 MW; 7162C9DB93BE0C7 CRC64;

Query Match 86.2%; Score 282; DB 2; Length 3010;
Best Local Similarity 85.7%; Pred. No. 4.8e-23;
Matches 54; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 ALENVIVNAAASAGTHGILWTFPCAAWYKGRVPGATYSILGLWPLLLLLALPPR 60
DB 747 ALENVIVNAAASVAGSHGILWTFPCAAWYKGRVPGAAVALYGVWPLLLLLALPPR 806
QY 61 AYA 63
DB 807 AYA 809

RESULT 8
Q90IX3 PRELIMINARY; PRT; 3010 AA.
AC Q90IX3;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MD9-2;
RX MEDLINE=20013335; PubMed=10544098; DOI=10.1006/viro.1999.9924;
RA Nagayama K., Kurosaki M., Enomoto N., Maekawa S.Y., Miyasaka Y.,
Tazawa J.I., Izumi N., Maruno F., Sato C.;
RT "Time-related changes in full-length hepatitis C virus and hepatitis
activity.";
RL Virology 263:244-253(1999).
DR EMBL; AF165062; AAD56197.1; -.
DR PIR; A61196; A61196.

DR PIR; PS0329; PS0329.
DR HSSP; Q8UJSL; ICWX.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR000345; C:ycb_heme_BS.
DR InterPro; IPR01545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR00745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_NS5a.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_core; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; Viral_RdRp; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
SQ SEQUENCE 3010 AA; 327254 MW; 9F1B0B3F536774FA CRC64;

Query Match 86.2%; Score 282; DB 2; Length 3010;
Best Local Similarity 85.7%; Pred. No. 4.8e-23;
Matches 54; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ALENVIVNAAASAGTHGILWTFPCAAWYKGRVPGATYSILGLWPLLLLLALPPR 60
DB 747 ALENVIVNAAASVAGSHGILWTFPCAAWYKGRVPGAAVALYGVWPLLLLLALPPR 806
QY 61 AYA 63
DB 807 AYA 809

RESULT 9
Q98UN3 PRELIMINARY; PRT; 562 AA.
AC Q98UN3;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.

```

OX NCBI_TaxId=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20499063; PubMed=11044085;
RA DOI=10.1128/JVI.74.22.10407-10416.2000;
RA Hadlock K.G., Lamford R.E., Perkins S., Rowe J., Yang Q., Levy S.,
RA Pilleri P., Abtignand S., Fong S.K.;
RT "Human monoclonal antibodies that inhibit binding of hepatitis C virus
RT E2 protein to CD81 and recognize conserved conformational epitopes.";
RU J. Virol. 74:10407-10416(2000).
[2]
RN SEQUENCE FROM N.A.
RP Chen M.K.C., Hadlock K.G., Yang Q., Chan L.;
RL Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF348705; AAK32686.1; -.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0019031; C:viral envelope; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR002519; HCV env.
DR InterPro: IPR002518; Pept_U39_HCV_NS2.
DR Pfam: PF01539; HCV_env; 1.
DR Pfam: PF01560; HCV_NS1; 1.
DR Pfam: PF01538; HCV_NS2; 1.
DR Coar protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polypeptide; Transmembrane.
FT NON_TER 1 1
FT CHAIN <1 39 E1.
FT CHAIN 40 402 E2.
FT CHAIN 403 >562 p7.
FT NON_TER 562 562
SQ SEQUENCE 562 AA; 61652 MW; 9C469E9F9CCFA9DA CRC64;

Query Match 85.9%; Score 281; DB 2; Length 562;
Best Local Similarity 85.7%; Pred. No. 1.4e-23;
Matches 54; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 1 ALENLVNMAASAGTHGILWFLVPCAMWYKGRVPGATYSLLGLMPLLLLLALPQR 60
DB 403 ALENLVNMAASVAGAHGILSLVFPFCAMWYIKGRVPGAAVALYGWVPLLLLLALPQR 462
OY 61 AYA 63
DB 463 AYA 465

RESULT 10
O80P91 PRELIMINARY; PRT; 1117 AA.
ID O80P91;
AC O80P91;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepcivirinae.
OX NCBI_TaxId=11103;
RP SEQUENCE FROM N.A.
RX MEDLINE=21904745; PubMed=11907242;
RA DOI=10.1128/JVI.76.8.4034-4043.2002;
RA Kalinina O., Norder H., Mukomolov S., Magnus L.O.;
RT "A natural intergenotypic recombinant of hepatitis C virus identified
RT in St. Petersburg.";
RU J. Virol. 76:4034-4043(2002).
DR EMBL: AY070174; AAL50213.1; -.
DR HSBP: G91R04; 1B77.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0019031; C:viral envelope; IEA.

```

```

DR GO: GO:0008236; F:serine-type peptidase activity; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR GO: GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO: GO:0019087; P:viral transformation; IEA.
DR InterPro: IPR002522; HCV_capsid.
DR InterPro: IPR002519; HCV_core.
DR InterPro: IPR002519; HCV env.
DR InterPro: IPR004109; Peptidase_S29.
DR InterPro: IPR002518; Pept_U39_HCV_NS2.
DR Pfam: PF01543; HCV_capsid; 1.
DR Pfam: PF01542; HCV_core; 1.
DR Pfam: PF01539; HCV_env; 1.
DR Pfam: PF01560; HCV_NS1; 1.
DR Pfam: PF01538; HCV_NS2; 1.
DR Pfam: PF02907; HCV_NS3; 1.
DR Coar protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polypeptide; Transmembrane.
FT NON_TER 1117 1117
SQ SEQUENCE 1117 AA; 121723 MW; 892B814C073656A8 CRC64;

Query Match 85.9%; Score 281; DB 2; Length 1117;
Best Local Similarity 85.7%; Pred. No. 2.6e-23;
Matches 54; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 1 ALENLVNMAASAGTHGILWFLVPCAMWYKGRVPGATYSLLGLMPLLLLLALPQR 60
DB 747 ALENLVNMAASVAGAHGILSLVFPFCAMWYIKGRVPGAAVALYGWVPLLLLLALPQR 806
OY 61 AYA 63
DB 807 AYA 809

RESULT 11
O68826 PRELIMINARY; PRT; 3010 AA.
ID O68826;
AC O68826;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polypeptide.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepcivirinae.
OX NCBI_TaxId=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=J33;
RA Cho M.J.;
RL Submitted (SEP-1991) to the EMBL/Genbank/DBJ databases.
DR EMBL: D14484; BAA03375.1; -.
DR PIR: A61196; A61196.
DR PIR: P00246; P00246.
DR PIR: P00804; P00804.
DR PIR: P80329; P80329.
DR HSBP: O8JYS1; 1CWK.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0019031; C:viral envelope; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO: GO:0003722; F:RNA binding; IEA.
DR GO: GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO: GO:0008236; F:serine-type peptidase activity; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO: GO:0006350; P:transcription; IEA.
DR GO: GO:0019079; P:viral genome replication; IEA.
DR GO: GO:0019087; P:viral transformation; IEA.
DR InterPro: IPR000345; CycC_heme_BS.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR011545; DEAD/DEAH_N.

```

```

DR InterPro: IPR002522; HCV_capsid.
DR InterPro: IPR002521; HCV_core.
DR InterPro: IPR002519; HCV_env.
DR InterPro: IPR002531; HCV_NS1.
DR InterPro: IPR000745; HCV_NS4a.
DR InterPro: IPR001490; HCV_NS4b.
DR InterPro: IPR002868; HCV_NS5a.
DR InterPro: IPR002166; HCV_NS5b.
DR InterPro: IPR001650; Helicase_C.
DR InterPro: IPR004109; Peptidase_S29.
DR InterPro: IPR009003; Pept_Ser_Cys.
DR InterPro: IPR002518; Pept_U39_HCV_NS2.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR Pfam: PF01543; HCV_capsid; 1.
DR Pfam: PF01542; HCV_core; 1.
DR Pfam: PF01539; HCV_env; 1.
DR Pfam: PF01560; HCV_NS1; 1.
DR Pfam: PF01538; HCV_NS2; 1.
DR Pfam: PF02907; HCV_NS3; 1.
DR Pfam: PF01006; HCV_NS4a; 1.
DR Pfam: PF01001; HCV_NS4b; 1.
DR Pfam: PF01506; HCV_NS5a; 1.
DR Pfam: PF00271; Helicase_C; 1.
DR Pfam: PF00998; Viral_RdRp; 1.
DR SMART: SM00487; DEXDC; 1.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
SQ SEQUENCE 3010 AA; 327151 MW; 7270F47984554FAD CRC64;

Query Match 85.9%; Score 281; DB 2; Length 3010;
Best Local Similarity 85.7%; Pred. No. 6.3e-23;
Matches 54; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ALENLVNNAASAGTHGILWLFVFCAMVYKGRVPGATSLGLWPLLLLLALPQR 60
DB 747 ALENLVNNAASVAGAHILSLVFPCCAMWIKGRVPGAAVALVGWPLLLLLALPQR 806

QY 61 AYA 63
DB 807 AYA 809

RESULT 12
O6GYR9 PRELIMINARY; PRT; 3010 AA.
ID O6GYR9;
AC O6GYR9;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_Taxid=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NS89;
RX PubMed=15218169; DOI=10.1099/vir.0.79984-0;
RA Kalinina O., Norder H., Magnus L.O.;
RT "Full-length open reading frame of a recombinant hepatitis C virus
strain from St Petersburg: proposed mechanism for its formation.";
RL J. Gen. Virol. 85:1853-1857(2004).
DR EMBL: AY587844; AAT00643.1; -.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0019031; C:viral envelope; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO: GO:0016787; F:hydrolyase activity; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003968; F:RNA-directed RNA polymerase activity; IEA.

```

```

DR GO: GO:0008236; F:serine-type peptidase activity; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR GO: GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO: GO:0006350; P:transcription; IEA.
DR GO: GO:0019079; P:viral genome replication; IEA.
DR GO: GO:0019087; P:viral transformation; IEA.
DR InterPro: IPR000345; CytC_heme_BS.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR011545; DEAD/DEAH_N.
DR InterPro: IPR002522; HCV_capsid.
DR InterPro: IPR002521; HCV_core.
DR InterPro: IPR002519; HCV_env.
DR InterPro: IPR002531; HCV_NS1.
DR InterPro: IPR000745; HCV_NS4a.
DR InterPro: IPR001490; HCV_NS4b.
DR InterPro: IPR002868; HCV_NS5a.
DR InterPro: IPR002166; HCV_NS5b.
DR InterPro: IPR001650; Helicase_C.
DR InterPro: IPR004109; Peptidase_S29.
DR InterPro: IPR009003; Pept_Ser_Cys.
DR InterPro: IPR002518; Pept_U39_HCV_NS2.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR Pfam: PF01543; HCV_capsid; 1.
DR Pfam: PF01542; HCV_core; 1.
DR Pfam: PF01539; HCV_env; 1.
DR Pfam: PF01560; HCV_NS1; 1.
DR Pfam: PF01538; HCV_NS2; 1.
DR Pfam: PF02907; HCV_NS3; 1.
DR Pfam: PF01006; HCV_NS4a; 1.
DR Pfam: PF01001; HCV_NS4b; 1.
DR Pfam: PF01506; HCV_NS5a; 1.
DR Pfam: PF00271; Helicase_C; 1.
DR Pfam: PF00998; Viral_RdRp; 1.
DR SMART: SM00487; DEXDC; 1.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
KW ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase;
KW Hydrolyase; Nonstructural protein; Polyprotein; Transmembrane.
SQ SEQUENCE 3010 AA; 326545 MW; D05AB2897F142A58 CRC64;

Query Match 85.9%; Score 281; DB 2; Length 3010;
Best Local Similarity 85.7%; Pred. No. 6.3e-23;
Matches 54; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ALENLVNNAASAGTHGILWLFVFCAMVYKGRVPGATSLGLWPLLLLLALPQR 60
DB 747 ALENLVNNAASVAGAHILSLVFPCCAMWIKGRVPGAAVALVGWPLLLLLALPQR 806

QY 61 AYA 63
DB 807 AYA 809

RESULT 13
Q9DTE1 PRELIMINARY; PRT; 3010 AA.
ID Q9DTE1;
AC Q9DTE1;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_Taxid=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUES=Serum;
RA Takahashi K., Iwata K., Matsumoto M., Matsumoto H., Nakao K.,
RA Hachihara T., Ohta Y., Kanai K., Maruo H., Baba K., Hijioka M.,
RA Mishiro S.;
RT "Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients

```

RT with hepatocellular carcinoma: the 'progression score' revisited."
RL Hepatol. Res. 20:161-171(2001).
DR EMBL/ AB049096; BAB1809.1; -
DR PIR/ A61196; A61196.
DR PIR/ P00246; P00246.
DR PIR/ P00252; P00252.
DR PIR/ P00253; P00253.
DR PIR/ P00254; P00254.
DR PIR/ P00804; P00804.
DR HSSP/ Q8JYS1; 1CWX.
DR GO/ GO:0016021; C:Integral to membrane; IEA.
DR GO/ GO:0019028; C:Viral capsid; IEA.
DR GO/ GO:0019031; C:Viral envelope; IEA.
DR GO/ GO:0005524; F:ATP binding; IEA.
DR GO/ GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO/ GO:0003723; F:RNA binding; IEA.
DR GO/ GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO/ GO:0008236; F:serine-type peptidase activity; IEA.
DR GO/ GO:0005198; F:structural molecule activity; IEA.
DR GO/ GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO/ GO:0006350; P:transcription; IEA.
DR GO/ GO:0019079; P:viral genome replication; IEA.
DR GO/ GO:0019087; P:viral transformation; IEA.
DR InterPro/ IPR000345; Cytochrome_B5.
DR InterPro/ IPR011545; DEAD/DEAH_N.
DR InterPro/ IPR002522; HCV_core.
DR InterPro/ IPR002521; HCV_core.
DR InterPro/ IPR002519; HCV_env.
DR InterPro/ IPR002531; HCV_NS1.
DR InterPro/ IPR000745; HCV_NS4a.
DR InterPro/ IPR001490; HCV_NS4b.
DR InterPro/ IPR002868; HCV_NS5a.
DR InterPro/ IPR002166; HCV_NS5b.
DR InterPro/ IPR001650; Helicase_C.
DR InterPro/ IPR004109; Peptide_S29.
DR InterPro/ IPR009003; Peptide_Ser_Cys.
DR InterPro/ IPR002518; Peptide_U39_HCV_NS2.
DR InterPro/ IPR007095; RNA_pol_DS_PS.
DR InterPro/ IPR007094; RNA_pol_PSVir.
DR Pfam/ PF01543; HCV_capsid; 1.
DR Pfam/ PF01539; HCV_env; 1.
DR Pfam/ PF01560; HCV_NS1; 1.
DR Pfam/ PF01538; HCV_NS2; 1.
DR Pfam/ PF02907; HCV_NS3; 1.
DR Pfam/ PF01006; HCV_NS4a; 1.
DR Pfam/ PF01001; HCV_NS4b; 1.
DR Pfam/ PF01506; HCV_NS5a; 1.
DR Pfam/ PF00271; Helicase_C; 1.
DR Pfam/ PF00998; Viral_RdRP; 1.
DR SMART/ SM00487; DEXDC; 1.
DR PROSITE/ PS00190; CYTOCHROME_C; UNKNOWN_1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
SQ SEQUENCE 3010 AA; 326845 MW; 90457AC819A32150 CRC64;
Query Match 85.9%; Score 281; DB 2; Length 3010;
Best Local Similarity 85.7%; Pred. No. 6.3e-23;
Matches 54; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
QY 1 ALENIYVINAASAGTGILMFLVPCAAWYVKGRLVPGATYSILGLMPILLILLALPQR 60
DB 747 ALENIYVINAASAGTGILMFLVPCAAWYVKGRLVPGATYSILGLMPILLILLALPQR 806
QY 61 AYA 63
DB 807 AYA 809

AC Q9J3G3;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DT Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_Taxid=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MD29;
RA Nagayama K., Kurosaki M., Enomoto N., Miyaoka Y., Marumo F., Sato C.;
RL Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL/ AF207770; AAF65960.1; -
DR PIR/ A61196; A61196.
DR PIR/ P00246; P00246.
DR PIR/ P00252; P00252.
DR HSSP/ Q8JYS1; 1CWX.
DR GO/ GO:0016021; C:Integral to membrane; IEA.
DR GO/ GO:0019028; C:Viral capsid; IEA.
DR GO/ GO:0019031; C:Viral envelope; IEA.
DR GO/ GO:0005524; F:ATP binding; IEA.
DR GO/ GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO/ GO:0003723; F:RNA binding; IEA.
DR GO/ GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO/ GO:0008236; F:serine-type peptidase activity; IEA.
DR GO/ GO:0005198; F:structural molecule activity; IEA.
DR GO/ GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO/ GO:0006350; P:transcription; IEA.
DR GO/ GO:0019079; P:viral genome replication; IEA.
DR GO/ GO:0019087; P:viral transformation; IEA.
DR InterPro/ IPR000345; Cytochrome_B5.
DR InterPro/ IPR001410; DEAD.
DR InterPro/ IPR011545; DEAD/DEAH_N.
DR InterPro/ IPR002522; HCV_capsid.
DR InterPro/ IPR002521; HCV_core.
DR InterPro/ IPR002519; HCV_env.
DR InterPro/ IPR002531; HCV_NS1.
DR InterPro/ IPR000745; HCV_NS4a.
DR InterPro/ IPR002868; HCV_NS5a.
DR InterPro/ IPR001490; HCV_NS4b.
DR InterPro/ IPR002166; HCV_NS5b.
DR InterPro/ IPR004109; Peptide_S29.
DR InterPro/ IPR009003; Peptide_Ser_Cys.
DR InterPro/ IPR002518; Peptide_U39_HCV_NS2.
DR InterPro/ IPR007095; RNA_pol_DS_PS.
DR InterPro/ IPR007094; RNA_pol_PSVir.
DR Pfam/ PF01543; HCV_capsid; 1.
DR Pfam/ PF01539; HCV_env; 1.
DR Pfam/ PF01560; HCV_NS1; 1.
DR Pfam/ PF01538; HCV_NS2; 1.
DR Pfam/ PF02907; HCV_NS3; 1.
DR Pfam/ PF01006; HCV_NS4a; 1.
DR Pfam/ PF01001; HCV_NS4b; 1.
DR Pfam/ PF01506; HCV_NS5a; 1.
DR Pfam/ PF00998; Viral_RdRP; 1.
DR SMART/ SM00487; DEXDC; 1.
DR PROSITE/ PS00190; CYTOCHROME_C; UNKNOWN_1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
SQ SEQUENCE 3010 AA; 327227 MW; 197835355650CACE3 CRC64;
Query Match 85.9%; Score 281; DB 2; Length 3010;
Best Local Similarity 85.7%; Pred. No. 6.3e-23;
Matches 54; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
QY 1 ALENIYVINAASAGTGILMFLVPCAAWYVKGRLVPGATYSILGLMPILLILLALPQR 60
DB 747 ALENIYVINAASAGTGILMFLVPCAAWYVKGRLVPGATYSILGLMPILLILLALPQR 806
QY 61 AYA 63

DB 807 AYA 809

RESULT 15

09J3G6 PRELIMINARY; PRT; 3010 AA.

AC 09J3G6 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Polypeptide.

OS Hepatitis C virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

OC Hepacivirus.

OX NCBI_TaxID=11103;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MD26;

RA Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;

RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF207767; AAF65957.1; -.

DR PIR; A61196; A61196.

DR PIR; P00246; P00246.

DR PIR; P00254; P00254.

DR PIR; P50329; P50329.

DR HSSP; Q8JYS1; ICWX.

DR MEROPS; S29.002; -.

DR GO; GO:0016021; C: integral to membrane; IEA.

DR GO; GO:0019028; C: viral capsid; IEA.

DR GO; GO:0019031; C: viral envelope; IEA.

DR GO; GO:0005524; F: ATP binding; IEA.

DR GO; GO:0008026; F: ATP-dependent helicase activity; IEA.

DR GO; GO:0003723; F: RNA binding; IEA.

DR GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.

DR GO; GO:0008236; F: serine-type peptidase activity; IEA.

DR GO; GO:0005198; F: structural molecule activity; IEA.

DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.

DR GO; GO:0006350; P: transcription; IEA.

DR GO; GO:0019079; P: viral genome replication; IEA.

DR GO; GO:0019087; P: viral transformation; IEA.

DR InterPro; IPR000345; CytC_heme_BS.

DR InterPro; IPR001410; DEAD.

DR InterPro; IPR011545; DEAD/DEAH_N.

DR InterPro; IPR002522; HCV_capsid.

DR InterPro; IPR002521; HCV_core.

DR InterPro; IPR002519; HCV_env.

DR InterPro; IPR000745; HCV_NS1.

DR InterPro; IPR001490; HCV_NS4a.

DR InterPro; IPR002868; HCV_NS5a.

DR InterPro; IPR002166; HCV_RdRp.

DR InterPro; IPR001650; Helicase_C.

DR InterPro; IPR004109; Peptidase_S29.

DR InterPro; IPR009003; Pept_Ser_Cys.

DR InterPro; IPR002518; Pept_U3_HCV_NS2.

DR InterPro; IPR007095; RNA_pol_DS_PS.

DR InterPro; IPR007094; RNA_pol_PSVir.

DR Pfam; PF01543; HCV_capsid; 1.

DR Pfam; PF01542; HCV_core; 1.

DR Pfam; PF01539; HCV_env; 1.

DR Pfam; PF01560; HCV_NS1; 1.

DR Pfam; PF01538; HCV_NS2; 1.

DR Pfam; PF02907; HCV_NS3; 1.

DR Pfam; PF01006; HCV_NS4a; 1.

DR Pfam; PF01001; HCV_NS4b; 1.

DR Pfam; PF01506; HCV_NS5a; 1.

DR Pfam; PF00271; Helicase_C1; 1.

DR Pfam; PF00998; Viral_RdRp; 1.

DR SMART; SM00487; DEXDC; 1.

DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN 1.

DR Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;

KW Polypeptidein; Transmembrane.

SQ SEQUENCE 3010 AA; 327166 MW; 74FAB6B80F24837B CRC64;

Query Match 85.9%; Score 281; DB 2; Length 3010;

Best Local Similarity 85.7%; Pred. No. 6.3e-23;

Matches 54; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ALENLVNNAASACTHGILMELVFPCAAMVKRGLVPGATYSILGLMPILLLLALPQR 60

DB 747 ALENLVNNAASVAGAHGISFLVFPCAAMYIKRGLVPGAAVALYGVMPILLLLALPPR 806

QY 61 AYA 63

DB 807 AYA 809

Search completed: October 4, 2005, 18:57:32

Job time : 65 secs